

[illegible]

LOCUS	AF282167	1829 bp	mRNA		PRI	30-AUG-2000
DEFINITION	Homo sapiens DRC3 mRNA, complete cds.					
ACCESSION	AF282167					
VERSION	AF282167.1	GI:9944939				
KEYWORDS						
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
AUTHORS	1 (bases 1 to 1829) Wu,K., Xu,Z., Wang,M., Xu,X., Han,Y., Cao,Y., Wang,R., Sun,Y. and Wu,M.					
TITLE	Cloning and expression analyses of down-regulated cDNA C6-2A in human esophageal cancer					
JOURNAL	Chung-Hua I Hsueh I Chuan Hsueh Tsa Chih 16 (5), 325-327 (1999)					
MEDLINE	99445952					
PUBMED	10514543					
REFERENCE	2 (bases 1 to 1829)					
AUTHORS	Wang,M., Chen,B., Wu,K., Xu,X., Han,Y., Cai,Y., Wang,J., Xu,Z. and Wu,M.					
TITLE	Direct Submission					
JOURNAL	Submitted (26-JUN-2000) National Laboratory of Molecular Oncology, Cancer Institute, CAMS, Panjiauyan, Chaoyang Qu, Beijing 100021, China					
FEATURES	Location/Qualifiers					
source	1..1829					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
	/chromosome="19"					
	/map="19q13"					
	13..1326					
	/note="tumor-associated protein"					
	/codon_start=1					
	/product="DRC3"					
	/protein_id="AA03038.1"					
	/db_xref="GI:9944940"					
CDS						
	/translation="MSPLSRPSPLPLARDVRLITGCPPLSACLYLAPRRHRRRL LPSGILLTLRKRPSEAEYDVLQIKYVALRDVNTLRNLGNLADSSPELLHFLGFLQIM IVNLSGILTRKRPSEAEYDVLQIKYVALRDVNTLRNLGNLADSSPELLHFLGFLQIM PYPYPERFGSEMPVPTDPOSRAWMDPEKQLQHRRRROYTQAQCRGMEVRRGRS AMPPLRLSYFLQSAPOVAVNGHRLDPESEPLLESTGKWLCAVDQARNRSL SVKRDVLEVDLSRKKMKVADPARGOGCYPXNLTPTPGRLHHSRPARLSNSTPP PPRAPAPALPRPRMDRPRMDRSCSLNGLDPSSEKTPRCCSSSTRRCRAAPRAA RDRASQSGPRPNRRSSARATPPRSAPGRPALAPGPTWMC"					
BASE COUNT	346 a 609 c 587 g 287 t					
ORIGIN						
	Query Match 85.8%; Score 1403.4; DB 89; Length 1829; Best Local Similarity 93.4%; Pred. No. 1.5e-242; Matches 1569; Conservative 0; Mismatches 9; Indels 102; Gaps 6;					
QY	35 gaggagctgtcgtacgcttcgagcgccgaagcgccctctgagggcgaggtacacgagctgtg 94					
DB	154 GAGGAGCTGTGCTGACCTGGGGCCAGCGCGCTCTGGAGGGAGCTACACGAGCTGCTG 213					
QY	.95 cagaagatcaaaatagcgttcacgcttcgctgtgcccgcgtgctgcggaacaatcgccgaccc 154					
DB	214 CAGAAGATCAAAATAGCGCTTCAGCGCTGCGCGCGCGGCGGAGCAACATCGCGACCCC 273					
QY	155 tccctccgagagctgttgcaactctcttcgaggtcctgtcagatgatgttgaaacagctcg 214					
DB	274 TCCTCTCCGCGAGACTGTTCACATTCCTTTTCCGGGCTCTGCAGATGATTTGTGAACACGTCG 333					
QY	215 gggggggccgagattgcgagagcagctgtgagggcgccgacatctgacatggatgacgtgagc 274					
DB	334 GGGGGGCGGAGATTGCGCGAGCACTGTGCGCGCGCGCGCATCTGACATCGAGATCCCTGTGGCG 393					
QY	275 ctgcgtcgagaaacagctcaactccagctgaaacagagctctgtgacctgcgttgaggacctg 334					
DB	394 CTGCTGCGGGGAGAACAGTCACTCTCAGCTGAACAGACGCTCTGACCTCGCTGGGGAGACTCG 453					
QY	335 tggaccgcgcccggcgtgagagctgtcccccggaggaaggagaccgccacatlagagaccgagttc 394					

```

Db 454 TGGACCGCGCGGCTGAGCTGTCCCGAGAGAGACCCCATACAGACCCAGATTG 513
OY 335 ttcaagcgcctggagacgcgcggttcaatgaccccgagagccgcgccttggagagccagtt 454
Db 514 TTCACGCGCTGGAGACCCCGCGGTACACCGACCCGAGAGCCGCGCTGGAGAGACCCAGTT 573
OY 455 gagaaacgcttcaacacagagcgagcgccg----- 487
Db 574 GAGAAACAGCTACAGCAGAGCGAGCGCCCGAGGTGACCCAGACGACACAGCAGGCG 633
OY 488 ----- 487
Db 634 CGAGGCTGGGAAGTCCGGGGGGCGCGCGTCCGCTGCGCCCGCTGACCCGACTGCT 693
OY 488 -----cagcaaaagcgcccgaggtcgctgtaatgtgtacagagagcttggagc 538
Db 694 TACTTCTTACAGCAAAAGCGCGCCCGAGGTGCTCAATGTGTCACCGAG--ACTTGGAGC 751
OY 539 cagaatctggaacctgaagcttgagtagagagagagaaatgggtctgtgttaattatg 598
Db 752 CAGAACTGAGCTCTAGCTGGAGTGAAGACAGCAAGAAATGGTCTCTGTGTAATTATG 811
OY 599 attccagcgcccgacacagcagtgagctgctggtcaagcagcgagagcttggagctcc 658
Db 812 ACTTCCAGGCGCGCAACAGAGTGAAGTGTGCTGCTCAACAGCGGAGCTACTGAGAGTCC 871
OY 659 tggatgacagtcgttaagtgtgtggaaggttcggagccagcgagggagagagatagtgc 718
Db 872 TGGATGAGAGTGTGAAGTGTGGAAGTTCGGGAGCCAGCGGGGAGAGAGGATATGTC 931
OY 719 cctcaacaatctgaaccttaaccttaacctcgagcgtcacacaagcgcaaaagccctgccc 778
Db 932 CCTTACACATCTCTGACACCTTACCCCGGAGCCCGGCTGACACAGCCCAAGCCCTTGCC 991
OY 779 gcaagcctgaagagactctctctcaacacacagcccgcccgcccgccacactcaagtc 838
Db 992 GCAGCCTGAAACAGCAGCTCTCTCTCACACACAGCCCGCGCCCGCCACTCTCAGTTC 1051
OY 839 tggcttggcccgcttggagacagcgcccgcttggagacgttcgata-nctcaagcgcttg 897
Db 1052 TGGCTCGCGCCCGCTGGGACAGCGCCCGCTGGGACGTGCTCAAGCTTCAAGCGCTTGG 1111
OY 898 accccagcgaagagagaaatctccacagatgcatatgltcaacgaagaaacttgagggc 957
Db 1112 ACCCCAGCAGAAAGAGAAATTT-TCGCAGATGCTCATGTGTAACGAGAACTGACAGCGCC 1170
OY 958 gcttggcccaagcgcttgcggagacagcgcgagtlccacagggcccgcccgcccgagac 1017
Db 1171 GCTTGGCCCAAGGCGCTCGGGACCGAGCCCGGAGTCCAGGGCCCGCGCCCGGAGAC 1230
OY 1018 cgcagctcagaccccggtctcgagcgtcccgantgcgcgcgtgctgagagcaagggt 1077
Db 1231 CGCAGCTCAGCCCGGCTCGGACGCTCCGAGGTCCGCGCTGCTGAGCCCAAGGCGCT 1290
OY 1078 ttaagctcgagacccgtgagacgcgttgggtgtgtcgtacccggggcgca-cttttcgcgtgc 1136
Db 1291 TTACTCTCCGGGACCGTGGACCGCTGGGTGTGCTGACCGGGGCCAGCTTTTCTCGCTGC 1330
OY 1137 agaagagagagcttgcggcggttgaagcccgagagaggggcaagctgtgtlacagccaggtca 1196
Db 1351 AGAAGGAGGAGCTGCGGGCGGTGAGCCCGGAGGGGGGCGTGTGTACAGCCAGGTCA 1410
OY 1197 ccttgcagcgcttgcgcgttggagagaaagagaaatgttcaagagcttggagagagttatg 1256
Db 1411 CCGTGCAGCGC-TCGCTGCTGAGAGCAAAAGAAAGTGTAGAGCTGAGAGCAGTATG 1469
OY 1257 gagaagcaaaagaaagagtggaagcgagtggaatgtgaagtgatcttgaacctgcagcag 1316
Db 1470 GAGAGCAAAAGAAAGAGTGTGAAGCGGAGAGGTGAAATGCAAGCTTATGACTTGCAGG 1539
OY 1317 cgccttcgcaaaagtgacagagcgccgttggagagacggactctcagactctcccaa 1376

```

```

Db 1530 GCGCTTGCANAAGATGACAGAGCCCGGTGGAGAACGAGACTCTCAGACTTCCCAA 1589
OY 1377 tagcgaagctgcatcttctgaagagatggccaatcgtccgcgccttcttccccatc 1436
Db 1590 TAGCGCAAGTGTGATCTTGAAGGATGGCCAACTGCTCCGCGCTTGCTTCCCCATTC 1649
OY 1437 ccggttgaacagacttaacgacttctgtcgaagctccctcccgagagagatctgagctg 1486
Db 1650 CCGGTGACAGACTTAAAGATCTTGTGCAAGTCCCTCCGAGAGAGATCTGAGTGCCTG 1709
OY 1497 ggaatggagagcgcttgaagacagcttcaagaaagcgcttaacagaccccgagagagtg 1556
Db 1710 GAGTGGGAGAGGCGGTGAGACAGTCTACGGAAGCCGTACAGACCCCGAGAGGAGTGC 1769
OY 1557 agtgaagcccttgaagcattgttaatactgcggccagcctaataacagccttgcgttaca 1616
Db 1770 AATGAGCCCTGAGCATTTGTAATATGCGGCCAGCGCTTAAACAGCCTCGCTTACCA 1829

RESULT 3
BC004907 2169 bp mRNA PRI 23-MAR-2001
LOCUS
DEFINITION Homo sapiens, similar to RIKEN cDNA 4632407K17 gene, clone
ACCESSION BC004907
VERSION BC004907.1 GI:13436193
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2169)
REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (21-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rudin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbiology.org
contact: amadan@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kettelman and Anuradha Madan
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov
Series: IRAL Plate: 11 Row: g Column: 15.
FEATURES
source
1. 2169
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:4642 IMAGE:3529476"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_lib="NIH MGC 17"
/lab_host="DH10B-R"
/note="Vector: pOT7"
/codon_start=1
/product="Similar to RIKEN cDNA 4632407K17 gene"
/protein_id="AAH04907.1"
/db_xref="GI:13436194"
/translacion="MDSGRVAAQEMLLRVSPDHVTLDPASKKELESYPLCAIYACD
AVMPGRSRSLLLVQEPERAPDVHFGQRLGAEILREDIOGALHNYRSGRGR
AAALRATQELQDRSPDAETPLQDRPRAVITSVVERGAGRGROAKPIDEAEAR
REPVTSSNADSPDLRGRGDLAVLQAEVEVDLNIHVPDVSFVRLQKSAEA
RVLEHRRRRRRRAGBGLTLTRAKPSEAEYDVLQIKITATSLARLNGNTADP

```


OY 1233 ggtcagagctggaagcagtgatggaagagcaaaagaaggtggaagcaggtggaag 1292
 DB 5603 GTGTGAGAGCTGAGAGCAGTATGATGAGAGCAAAAGAGAGGTGAGAGGAGGAGAA 5662
 OY 1233 atggaggtcatttgacccgcgcgcgccttcgcaagaagtgacgaagcccggtggaga 1352
 DB 5663 ATGGAGGTCTATTGACCTGCGAGGCGCCCTTGCAAAAGTACAGAGGCCCGTGGAGAA 5722
 OY 1353 agggactctcgaactctccccaatgaggaatgcatctcttgaaagatgccaatctg 1412
 DB 5723 ACGGACTCTCAGACTCTCCCAATAGCCGAGAGTCGATCTTGAAAGAGTGCCAAATCTG 5782
 OY 1413 ctccagcctcgtctctcccccaccccggtgagacagactaacatcctctgcagctccc 1472
 DB 5783 CTCCGCGCCCTGCTCTCCCGCATCCCGGTGAGACACACTTAACGATCCTTGCTGCACTCC 5842
 OY 1473 tccggaagagatctggaactgctggaagtggagagggcgtggaagacagctctcgaagaac 1532
 DB 5843 TCCGAGAGGATCTGACTGCTGGAGAGTGAGGCGCTGAGACAGTCTACGGAAGAC 5902
 OY 1533 gctagcagaccccgagaggggtgagctggaagcctgagcatgtatgagggccaagcc 1592
 DB 5903 GCTTACGACACCCCGAGAGGGTGCAGTGAGAGCCCTGAGCATTGTATATGCGGCCAGCC 5962
 OY 1593 tataaacagcctcgcgtgcttagcagatggtgtccant 1631
 DB 5963 TATAAACAGCCCTCGCTGCTTACGATGCTGTGCCAAGT 6001
 RESULT 5
 AC019238/c AC019238 204340 bp DNA HTG 17-AUG-2000
 LOCUS Homo sapiens chromosome 19 clone RP11-700B5, WORKING DRAFT
 DEFINITION
 SEQUENCE, 12 unordered pieces.
 ACCESSION AC019238.5 GI:9838316
 VERSION AC019238.5
 KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 204340)
 AUTHORS Waterston,R.H.
 JOURNAL The sequence of Homo sapiens clone
 TITLE Unpublished
 REFERENCE 2 (bases 1 to 204340)
 AUTHORS Waterston,R.H.
 JOURNAL Direct Submission
 TITLE Submitted (30-DEC-1999) Genome Sequencing Center, Washington
 JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 COMMENT On Aug 17, 2000 this sequence version replaced gi:9280808.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 ----- Project Information -----
 Center project name: H.NH0700B05
 ----- Summary Statistics -----
 Sequencing vector: M13; 848
 Sequencing vector: plasmid; 16%
 Chemistry: Dye-primer ET; 84% of reads
 Chemistry: Dye-terminator Big Dye; 16% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 199676 bases at least Q40
 Consensus quality: 200928 bases at least Q30
 Consensus quality: 201749 bases at least Q20
 Insert size: 236000; agarose-fp
 Insert size: 204523; sum-of-contigs
 Quality coverage: 6.54 in Q20 bases; agarose-fp
 Quality coverage: 7.57 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 12 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 2752: contig of 2752 bp in length
 2753 2852: gap of unknown length
 2853 5901: contig of 3049 bp in length
 5902 6001: gap of unknown length
 6002 12965: contig of 6964 bp in length
 12966 13065: gap of unknown length
 13066 21018: contig of 7953 bp in length
 21019 21119: gap of unknown length
 21119 30399: contig of 9281 bp in length
 30400 40958: gap of unknown length
 40959 41059: contig of 10459 bp in length
 41059 52180: gap of unknown length
 52181 52280: contig of 11122 bp in length
 52280 65276: gap of unknown length
 65276 65375: contig of 12995 bp in length
 65375 95161: gap of unknown length
 95161 95261: gap of unknown length
 95261 122538: contig of 27277 bp in length
 122538 122539: gap of unknown length
 122539 155779: contig of 33140 bp in length
 155779 204340: gap of unknown length
 204340 48462: contig of 48462 bp in length.

FEATURES
 source 1. 204340
 /organism="Homo sapiens"
 /db_xref="Caxon:9606"
 /chromosome="19"
 /clone="RP11-700B5"
 1..2752
 /note="assembly_name:Contig16"
 2853..5901
 /note="assembly_name:Contig17"
 6002..12965
 /note="assembly_name:Contig18"
 13066..21018
 /note="assembly_name:Contig19"
 21119..30399
 /note="assembly_name:Contig20"
 30500..40958
 /note="assembly_name:Contig21"
 41059..52180
 /note="assembly_name:Contig22"
 52281..65275
 /note="assembly_name:Contig23"
 65376..95161
 /note="assembly_name:Contig24"
 95262..122538
 /note="assembly_name:Contig25"
 122639..155778
 /note="assembly_name:Contig26
 clone_end:77
 vector_side:right"
 155879..204340
 /note="assembly_name:Contig27
 clone_end:SP6
 vector_side:left"

BASE COUNT 48032 a 51557 c 52790 g 50855 t 1106 others
 ORIGIN
 Query Match 25.4%; Score 414.8; DB 65; Length 204340;
 Best Local Similarity 83.6%; Pred. No. 7.3e-66;
 Matches 534; Conservative 0; Mismatches 13; Indels 92; Gaps 2;

OY	1085	cggagccgtagacgcgtcggctgtagaccggaggca-cttttcgcgtcgaaagga	1143
Db	185794	CAGGACCGTGGACGCCCTGGGTGTCATCCGGGGCAGCTTTTCGTGGAGAGGA	185735
OY	1144	ggagctcgcyggygltgtagccccggaggaggccaagtgtctacaaggtcaacctga	1203
Db	185734	GGAAGTCGGGGGGGTGAGACCCCGAGAGGGGCACTGTGTACAGCCAGGTCAACGTGA	185675
OY	1204	ggcgct-----	1209
Db	185674	GGCGTCGCTGTGAGGTAGACGACCGCTGCTCCGTGGGTCTGGGTAGGTTGGGATG	185615
OY	1210	-----gcctgcgtggagcaagaagaaa	1232
Db	185614	ACGAGGCGCTCGGACGCCACGAGCAAACCATTTCCACCOCGCCCTCCAGACAAGAAGAAA	185555
OY	1233	gtctcagagcttgagagcagtgatgatgagaagcaaaaaagaagtgtagaagcgagtgtaa	1292
Db	185554	GTGTCAAGAGCTGGAGAGCGATGATGGAGAACAAAGAAGGTGGAGAAGCGAAGTGGA	185495
OY	1293	atggaggtcatttgaacctgccaagcgcccctcgcaaaagtlgaagagggcccgtyggaga	1352
Db	185494	ATGGAGGTCTATTGTGACTCCGCAGCGGCCCTTCGCAAAAGTGAAGAGGCCGCCCTGGAGAA	185435
OY	1353	acggagctcttaagactctcccccaatagcgaagatcgatctcttgaagatggccaatcg	1412
Db	185434	ACGGACTCTCTAGACTCTCCCCAATAAGCGAAGTGTGATTTCTGAAGAGAGGCAATCTG	185375
OY	1413	ctcgcgccttgctctcccccatacccggttgagacagacttaagatcatctgtcagttcc	1472
Db	185374	CTCCGGCCCTGGTGTCTCCCCATCCCGGTGACAGACTTAACATCTTGTCTCAATCC	185315
OY	1473	tccggagagagatctgagatcgtgagctgtggaggtggtggaagcgtggaacagcttacggaagc	1532
Db	185314	TCCGAGAGGAGATCTGTGACTGTGGAGTGGGGAGGCGCTGAGACAGTCAAGGAAAAAC	185255
OY	1533	gctagcagaacccccggagagggtgcaagtggagcccttgagctttaatatgagccccagcc	1592
Db	185254	GCTAGCACACCCCAGAGGGGTGAGTGGACCCCTGTGACATTTGAATATGCGGCCACAGCC	185195
OY	1593	tataaacagacctccgctgcttagcagatgagtgtgccantt	1631
Db	185194	TATAACAGCCTCCGTGCTTAGCAGATGAGTGTGCCAACT	185156
RESULT	6		
AF282168			
LOCUS	AF282168	5985 bp DNA	PRI 30-AUG-2000
DEFINITION	Homo sapiens DRC3 gene, complete cds.		
ACCESSION	AF282168		
VERSION	AF282168.1	GI:9944941	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 5985) Wu,K., Xu,Z., Wang,M., Xu,X., Han,Y., Cao,Y., Wang,R., Sun,Y. and Wu.M. Cloning and expression analyses of down-regulated cDNAs C6-2A in human esophageal cancer Chung-Hua I Hsueh I Chuan Hsueh Tsa Chih 16 (5), 325-327 (1999)		
AUTHORS			
TITLE			
JOURNAL			
MEDLINE	99445952		
PUBMED	10514543		
REFERENCE	2 (bases 1 to 5985)		
AUTHORS	Wang,M., Chen,B., Wu,K., Xu,X., Han,Y., Cai,Y., Wang,J., Xu,Z. and Wu.M. Direct Submission Submitted (26-JUN-2000) National Laboratory of Molecular Oncology Cancer Institute, CAMS, PUMC, Panjiayuan, Chaoyang Qu, Beijing 100021, China		
TITLE			
JOURNAL			
FEATURES	Location/Qualifiers		
SOURCE	1..5985		

Query Match	Best Local Similarity	83.8%	Score 405.6;	DB 89;	Length 5985;
Matches 523;	Conservative	0;	Mismatches 9;	Indels 92;	Gaps 2;
QY 1085	cgggacgctgagcgcgcgcgtggtgtgtcgaacggggcgca-ccttcctgcgtcagaaga	1143			
Db 5362	CAGGACCGGTGACGCGGTGGGTGCTCACCAGCGGCGGAGCTTTCTCGTCGACAAAGA	5421			
QY 1144	gaagctgcggcggtgagagcccgagaagggcgagctgtgtacagcaggtacacgtgtga	1203			
Db 5422	GGAGCTGGGGGGGGTGAAGCGCCCGAAGAGGGGGGACAGTGTGTACACGAGCTACCGTGCA	5481			
QY 1204	gcgcgtt-----	1209			
Db 5482	GGCGTCGCTGTGAGGTGAGCGGACCGACGCTGCTCCTGGGTGTGAGGTTGGCATG	5541			
QY 1210	-----gcgtgcgtggaagacaagaagaa	1232			
Db 5542	ACGAGGCGCTCGGACCCACAGGACAAAGCCATTTCACCCCGCCCTTCACGACAAAGAGAA	5601			
QY 1233	gtgtcagaagcttgagacagtgatgtgagaagacaagaagaaggttgaaagcgaggtgaa	1292			
Db 5602	GTGTGACAGACTGAGAGCAGTGTATGAGAAGCAAAAGAAAGAGTGTGAAGCGAGGTGAA	5661			
QY 1293	atggaggttcatttgaactgtgcacgaagcgcccttcgcaagaagtgaagagcccggtgaga	1352			
Db 5662	ATGGAGGTTCATTGACCTTCACGAGCGCCCTTCGCAAAAGATGACAGGCCCGCGTGGAGA	5721			
QY 1353	acggacgtcctcagactctccccaatagcgaagatgcactctctgaagatgtggcaatctg	1412			
Db 5722	ACGGACTCTCTAGACTCTCCCAATAGCGGAAGTGTGATCTTGTGAAGATGGCCCAATGTG	5781			
QY 1413	ctcggccctggtcttcccccaccccggttgacagacttaacgactccttgttcgaatccc	1472			
Db 5782	CTCCGCGCCCTGGTCTTCCCCCATCCCGGTGAGACAGACTTAAGATCTTGTCTGCAAGTCCC	5841			
QY 1473	tcggagaagagatctgagactgtgctggaatgtggagaggggttgagagacagttacaggaagc	1532			
Db 5842	TCCGAGAGAGATCTGCACTGGCTGGAGTGGGAGGCGCTGGAGACACTTACGAAAGC	5901			
QY 1533	gctaacgacaccccgagaaggggtgcagtgagagcccttgacatgtlaatatgcgcccagcc	1592			
Db 5902	GCTACGACACCCCGAGAGGGGTGCAATGAGACCCCTGACACTTGAATATGTGGCGCCAGCC	5961			
QY 1593	tataaacagcctcgtgtcttagca	1616			
Db 5962	TATAAACAGCCTCGGTGCTTAGCA	5985			

RESULT 7
AK025588
LOCUS AK025588 3102 bp mRNA PRI 29-SEP-2000
DEFINITION Homo sapiens cDNA: FLJ21935 fls, clone HEP04373.
ACCESSION AK025588
VERSION AK025588.1 GI:10438149
KEYWORDS oligo cloning; fls (full insert sequence).
SOURCE Homo sapiens hepatoma cell_line:hep02 cDNA to mRNA, clone_11b:HEP clone:HEP04373.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
Kawabata, A., Hiki, T., Kobatake, N., Inagaki, H., Ikema, Y., Shihabara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S., Shihabara, T., Tanaka, T., Ota, T., Suzuki, Y., Obayashi, M., Nishii, T., NEDO human cDNA sequencing project
2 (bases 1 to 3102)
Unpublished (2000)
Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishii, T., Isogai, T., Shihabara, T., Tanaka, T. and Nakamura, Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) to the DDBJ/EMBL/GenBank databases. Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@leims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: 5'-6' and one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).
FEATURES
source location/Qualifiers
1..3102
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HepG2"
/cell_type="hepatoma"
/clone="HEP04373"
/clone_11b="HEP"
/note="Cloning vector pME18SFL3"
82..2313
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB15180.1"
/db_xref="GI:10438150"
translation="MOVMDRHPOTGATLRSLPLATODTMSGASGCPGATNG SLGRSDGAKKSPKIDFEORRKYNSNYIMETSYHONHATPTMDSEATSDDA TRKIVOLSSKEKINTOEMLVNDOSLALDIESOELEDPPLTVONSQVYNDLRY PSLVLVCQDSSEKSPDVHFFHCDSEAEVLEHEDSLADCRKLRKRPOTLKHOE KIRKROSLPPPOGPAPIPQHRGSDSPBAENRQPOVPLSEPGRRRESDEPRALV AOKIEKTEQIINLCALDDIDEMFVARLQKAAEFKQONKKKKKKKAPAEGLTLRA RPPSEGFIDCFOKIKALINLAKIIONPSAEELVHFLFEGPLDLVNGVSGDIA RVSCEPLISRAVDLRLGHILPKEMSLMESIGESMRSEPRPREPOVPLVYPRKSG WEPPYDVLOEARPMVEGLASAPIEIVSYNSQSIINSKSHSTSEPTTPGDLRPVSS PTHRGVQTPPAMATYKILVDFTRKNRNEUSVLDEVELEDEDRQWKKLSRSGQA GYVPCNIGEARPEBDAGAPFBOAGOKYWPAPSPHKLPPSPGNDELMORDEVNDL LIRKISNIRAPORHFRIVERSOPASOPLTYSGPDEVRAMLEAKAFSPRIVENLGLT GPOLFSLNKEELKKYCGEEGVRYVQLTMQAFLEKKQSGSELELMKMFISMNRG EDS"

BASE COUNT 678 a 972 c 974 g 478 t
ORIGIN

Query Match 13.1%; Score 214.2; DB 89; Length 3102;
Best Local Similarity 54.3%; Pred. No. 3.6e-29;
Matches 698; Conservative 0; Mismatches 480; Indels 107; Gaps 9;

QY 25 ggcgcgtgaggagagcgttgctgacgcctgcgagcccaagccgcctcgaggccgagtaaac 84
DB 1050 ggcgcgcagagagagggcgcttcctcacaactgcggcagagggccctcctcgaggcgagttcat 1109
QY 85 cgactgctgcaagaatcaagtaacgcttcagcctgtctgcccgcgtgcgcgcaaac 144
DB 1110 cgactgcttcacgaaatcaacgcttcagcctgtctgcccgcgtgcgcgcaaac 1169
QY 145 ggcgcgcgcctcctcgcgagcctgttgcaacttccttcgcggccttcgagagatgct 204
DB 1170 ccagAACCCAGCGCGCGGAGACTGTCGACTTCTTCCGAGCTGAGACTGACTGCT 1229
QY 205 gaacacgtcggggggcgagctgcgcgagcagtgctgcggcgccgacatctgacatcga 264
DB 1230 caacacgtcgcagtgccgagcagacatgcacgctccgctccgacacgtcctccgaga 1289
QY 265 tgcgctgagcgtgcgtgcgggagcaacgtaactcaacgtgaagaacgagctctgactgct 324
DB 1290 tgcgctgagcgtgcgtgcgggagcaacgtaactcaacgtgaagaacgagctctgactgct 1349
QY 325 ggggagactcgtgagccgc 381
DB 1350 ggggagactcgtgagccgc 1409
QY 382 cagacacgagcttcctcagcgtgcgtgcggcgccggtcaactgaccccgagcgcgcgtc 441
DB 1410 cctgcgcagagcttcacagcgtgcgtgcggcgccggtcaactgaccccgagcgcgcgtc 1469
QY 442 ggaagagcc-----agtgagaagaacagctaacgacgag 475
DB 1470 ggaagagcc-----agtgagaagaacagctaacgacgag 1529
QY 476 cggagcgccgcgagcaagaagcgcccgagcgtgcgtgcgtgcgtgcgtgcgtgcgtgc 535
DB 1530 cggagcgccgcgagcaagaagcgcccgagcgtgcgtgcgtgcgtgcgtgcgtgcgtgc 1589
QY 536 agccagatctgagcctcagcctgagctgagctgagcagacga-----ggaac 578
DB 1590 agccagatctgagcctcagcctgagctgagctgagcagacga-----ggaac 1649
QY 579 atggctcctgtgtaattatgacttcgaagcccgcaacagcagtgagctgctgacgaac 638
DB 1650 atggctcctgtgtaattatgacttcgaagcccgcaacagcagtgagctgctgacgaac 1709
QY 639 ggggagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgag 698
DB 1710 ggggagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgag 1769
QY 699 ggggagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgag 758
DB 1770 ggggagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgag 1821
QY 759 ccaacgcaaaacccctgc 818
DB 1822 ccaacgcaaaacccctgc 1881
QY 819 cccgc 878
DB 1882 cccgc 1930
QY 879 cgatantcacaagcgttgagccccaagcgaagaggaatctcctccagatgctcatalgctc 938
DB 1931 cgatantcacaagcgttgagccccaagcgaagaggaatctcctccagatgctcatalgctc 1977
QY 939 aacgaggaactcgaagcgtgcgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 998
DB 1978 aacgaggaactcgaagcgtgcgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2013
QY 999 gggccgc 1058
DB 2014 gggccgc 2070
QY 1059 tggctgcaagcaagggcttagcttcgcgagcgtgagcgcgtggtgtgtgacgcgg 1118

Db 2071 TGGCTGGAAGCCAGAGCCCTTACGCCCGGATGTTGAAACATCTGGGACCTCGACCGGG 2130
 QY 1119 GCGCA-cttcttcgctgcaagaagagagagctgagcggtgagcccgagagagagagca 1177
 Db 2131 CCGACAGCTCTTCTCCCTCAACAAGAGAGAGCTGAAGAAAGTGTGGCGGAGAGGCGGTC 2190
 QY 1178 cgtgtgtacagcagagtcacgctgacgcttcgctgctgctgagagacaaagaaagtgctc 1237
 Db 2191 CCGGTGTACAGCAGCTCAGCATGAG-AAGGCTTCTGTGAGAGACAGCAAGTGGGTC 2249
 QY 1238 agagctgagagcagtgatgagaaag 1262
 Db 2250 GGAGCTGGAGAACTCATACACAG 2274
 RESULT 8
 BC002474 2646 bp mRNA PRI 16-MAR-2001
 LOCUS Homo sapiens, similar to hypothetical protein FLJ21935, clone
 DEFINITION MG:3088, mRNA, complete cds.
 ACCESSION BC002474
 VERSION BC002474.1 GI:12803312
 KEYWORDS MGC.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2646)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (05-FEB-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert_Strausberg@nih.gov
 Tissue Procurement: DCTD/BNP
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nih.gov
 Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
 Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Snyder, B., Stantirip, S., Thomas, P.J.,
 Tlonsong, E.E., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A.,
 Zhang, L.-H. and Green, E.D.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 5 Row: h Column: 6.
 Location/Qualifiers
 1. 2646
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MG:3088 IMAGE:3347644"
 /tissue_type="Kidney, renal cell adenocarcinoma"
 /clone_id="N1H_MGC_14"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 887..1870
 /codon_start=1
 /product="Similar to hypothetical protein FLJ21935"
 /protein_id="AAH02474.1"
 /db_xref="GI:12803313"
 /translation="MSLWSLGSWMRPRSEWPREPVLPYVFKFHSGMPEPVDVIOE

APWEVGLASAPIEVSFVSNOSIRNSOKHSPTSEPPPGDAPLPVSSPHTRGCOPT
 PAMAKYVILYDFARANANLSELYKDELEVLEEDGROMMKLSRSGOAGVPCNILE
 ARPDAGAPFEQACOKYWGSPSPHKLPSPSPGKMDKDELMOHDEVNELIRKISNIRA
 QPOHPRVEREOPVSOPLTVSSGDEVRAMLEAKAFSPRIYENGILITGQPSLANE
 ELKRVCEEGEGRVYSQILTMOKAFLEKQSSSELELNMKHSNMORGEOS"
 BASE COUNT 574 a 829 c 825 g 418 t
 ORIGIN
 Query Match 13.1%; Score 213.6; DB 91; Length 2646;
 Best Local Similarity 53.8%; Pred. No. 4.8e-29;
 Matches 691; Conservative 0; Mismatches 487; Indels 107; Gaps 8;
 QY 25 ggcgctgaggagagctgctgacgctgagccagcccgctgagggcgagctaac 84
 Db 607 GGGCCAGACAGAGGGCGCTCTCACACTGCGGGACGGCCCTCTGTGAGGGCGACTTAT 666
 QY 85 cgaagctgctgcaagaatcaagctaacgcttgcctgagcccgctgagcgagcaaat 144
 Db 667 CGACTGCTTCCAGAAATCAAGCTGGCATTAACCTGTGGCAAAAGTGCAGAAAGACAT 726
 QY 145 ggcgagccctctctcggagctgttgacttcttcttgagctctgagagatgtct 204
 Db 727 CCAGAACCCAGCGCGCGAGCTGTGCTCTTCTGAGGCTGTGAGCTGTATGCT 786
 QY 205 gaacacgctcggaggcgagctgagctgagctgagcgagctgagctgagctgag 264
 Db 787 CAACACCTGCGAGTGCGCCAGACATCGACGCTCGCTCCGCCACGCTCTCCCGAGA 846
 QY 265 tgcgctgagctgctgctgaggaacagctcaacgctgaaacagagcttgagctgct 324
 Db 847 TGCGGTGACTTCTGCGCGCGCGCGCTGCTTCTTGAAGAGATGTGCTGTGGAGTCACT 906
 QY 325 ggggagctcggagcccgcccgagctgagcttcccgagag---gagggagcccata 381
 Db 907 GGGAGAGAGCTGGATGCGGCGCGCTTCCGAGTGGCGCGGAGACAGGCTGCTCTTA 966
 QY 382 cagacagagcttctcagagctgagagcgagctgagctgagcagcgagcgagctg 441
 Db 967 CGTGCCCAAGTTCCACACAGCGGTGGAGCTCTGTGATGCTGTCGAGAGGCGCCCTG 1026
 QY 442 gggagagcc-----agttgaaacagctacagcagag 475
 Db 1027 GGAGGTGAGGGGCTGCGCTGCGCCCATGAGAGAGGTGAGTCCAGTACCCACAGATC 1086
 QY 476 cggagcgcccgagcaaaagcgcccgagctgctgctgcaatgctacccgagagctgg 535
 Db 1087 CATAGAAATCTCCAGAAAGCACAGCCCATCTTCAAGAGCCACCCCGGGGATGCCCT 1146
 QY 536 agcagaatctgagctcagctgagctgagctgagagagagag-----aaa 578
 Db 1147 ACCACAGTACAGTCCACACTACTACAGAGGGGTGCGACGACACACAGCATGGCCAA 1206
 QY 579 atggctctgtgtaatatgacttccagagcccgagcaagcagtgagctgtgctgaaga 638
 Db 1207 GTACGTCAAGATCTGTATGATCTTACAGCGCCGAAATGCCAAGCATATGCGCTCAA 1266
 QY 639 ggcgagctgagagctgagctgagctgagctgagctgagctgagctgagctgagc 698
 Db 1267 GGATGAGGTCTCTAGAGTGTCTGAGAGACAGCGCGGATGTGTGAGAGCTGCGACGCCAG 1326
 QY 699 gggcagagagagatgctgcttcaaacatctgacacccctaccccgagcccgagctgca 758
 Db 1327 CGGCGAGGCGGGGTACGTGCGCTTCAACAT-----CTTAGCGAGGGCGGACCGGAG 1378
 QY 759 ccaagcacaagagccctgagcagcagctgagcagcagcttctctcaccacacagccagc 818
 Db 1379 GAGCGCGGCGCGCTTCTGAGAGCGCGGTCAAGATCTGGGGCGCGCGCCCGCCACACC 1438
 QY 819 ccggagccacacacagctgagctgagcccgctgagcagcgcccgctgagcagctg 878
 Db 1439 CACAAGCTACCCCAACAGCTT-----CCGGGG 1465


```
Db 1524 ATTGACACACGAGGAGCTCCACATGCTAAACAGGGGCTACACCAACACAGCCAGG 1583
Oy 622 gggcgtcgtcgaacagcggagcgtactgtaggtccctgagatcagctgtaagtgtg 681
Db 1584 ACCAAGTACGTAAACATCTCTATATGATTTACACAGCGCCGAACGCCAATGACTATCTGTG 1643
Oy 682 ---aaagttcggagaccagcgggagagagatatgtccctacaacactc--gacac 736
Db 1644 CTCAGAGAGAGAGAGTCTTAGAGAGTGTGGAGAGAGCGCCGCACTGGTGAAGCTACGAAT 1703
Oy 737 cctaccgggagcccggtgtgacacagccaaagccctgcccagcctgtgaacgaactc 796
Db 1704 CCCAGTGGCCAGCGCGGCTATGTGCCCTGCACATCTGGCTGAAGCTGGCAGGAGGAC 1763
Oy 797 cctccacacacagaccccccagcccgcccccctccacactctgtgctgcccgcctgtg 856
Db 1764 GTGGGTGCCCTTTGGACAGTCTGGCCAGAAATCTGGGGTCCAGCCAGCCCAACCCAC 1823
Oy 857 acagggcccgctggagacagctgcatanctcaacgct--tggaccacagcagagaagag 914
Db 1824 AAACGTGCCCCCAATTTTGCAGAGGAACAAGAGAGCTAATCCACACATGGACGAGGTC 1883
Oy 915 aaattctccagatgtctcatcgtcaagaggaactgagcgccctggccagagggcg 974
Db 1884 AATGACGAGCTCATGAAAAAGATCAAGCCACATCAAGACACAGCCGCGCAACTTCCGC 1943
Oy 975 tgggagcagagcggcagctcccaaggcccgcccgcccggaacgagctcagccggcg 1034
Db 1944 GTGGAGCGGACGACGCACT-----GCACCTGCGCCCTTACCTTCGAG 1985
Oy 1035 tggagcgcctccgagngtccgcctgctgctcagagncaaagagctttagctccggagacgt 1094
Db 1986 TCAGGTCCGGAGTAGAGTCCGCCCTGTGGAGAGGCCCTTCACTGATCCAGGATCGTG 2045
Oy 1095 gacgcgtgtgtgtgtgtgacgcggggcgca--ctttctgctgctcaagagagagctgg 1153
Db 2046 GAGAACCTGGCATCTTAACCTGACCCCGACGCTCTTCCCTCAACAGAGAGAACTGAAA 2105
Oy 1154 ggggtgagcccgagagaggggagcgtgtgtcaagcaggtacacgtgagcgcttcgct 1213
Db 2106 AAAGCTGCGCGGAGAGAGGCGAGTCCGTGTACAGCCACCTCACGCTTCAG--AAGCCCTT 2164
Oy 1214 gcttgagagcaaaagagaaagtgtcagagcttggagcgatgagagaaagcaaaagaa 1273
Db 2165 CTTGGAGAAACAACAAGAGGGGTCAAGACTGGAGAACTCATGAGCAAGATCAGAGGGC 2224
Oy 1274 ggtggaagcgca 1285
Db 2225 CGAGGACACCTA 2236

RESULT 11
AC011476
LOCUS AC011476 188302 bp DNA HTG 18-JUL-2000
DEFINITION Homo sapiens chromosome 19 clone CTC-550B14, WORKING DRAFT
SEQUENCE AC011476
AC011476.6 GI:9256296
ACCESSION AC011476.6
KEYWORDS HTG; HTGS_PHA5E2; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 188302)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 19
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 188302)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Sublission
JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
```

```
COMMENT
On Jul 18, 2000 this sequence version replaced gi:8576068.
-----Genome Center
Center: Joint Genome Institute
Center code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 396418, BC361442
Center clone name: CIT-HSPC_550B14
-----
Summary Statistics
Consensus quality: 185362 bases at least Q40
Consensus quality: 186527 bases at least Q30
Consensus quality: 186988 bases at least Q20
Estimated insert size: 189730; agarose-IP estimation
Estimated insert size: 187852; sum-of-ctdigs estimation
Quality coverage: 8.26 in Q20 bases; agarose-IP estimation
Quality coverage: 8.36 in Q20 bases; sum-of-ctdigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
1 19837: contig of 18837 bp in length
19838 19937: gap of unknown length in length
19938 66572: contig of 46635 bp in length
66573 66673: gap of unknown length
66673 69024: contig of 2352 bp in length
69025 69125: gap of unknown length
69125 81795: contig of 12670 bp in length
81795 81895: gap of unknown length
81895 88331: contig of 6437 bp in length
88331 88332: gap of unknown length
88332 94151: contig of 5720 bp in length
94151 94152: gap of unknown length
94152 147534: contig of 53283 bp in length
147534 147534: gap of unknown length
147535 147635: gap of 6562 bp in length
147635 154196: contig of 6562 bp in length
154196 154296: gap of unknown length
154296 166008: contig of 11712 bp in length
166008 166109: gap of unknown length
166109 171710: contig of 5602 bp in length
171710 171711: gap of unknown length
171711 171810: gap of unknown length
171810 188302: contig of 16492 bp in length.
*
Location/Qualifiers
1. 188302
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTC-550B14"
/clone_11b="CalTech human BAC library C"
BASE COUNT 47414 a 45925 c 47597 g 46366 t 1000 others
ORIGIN
Query Match 9.2%; Score 151; DB 62; Length 188302;
Best Local Similarity 100.0%; Pred. NO. 1.7e-18;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 199 gatttggaacacgttcgggggggagcgttcgagacagtgtagggcgccgcatctgac 258
Db 186443 GATTGTGAACACGTCGCGGGGGCCGAGTTTCGGACACAGTGTGCGGCCCGCATCTGAC 186502
Oy 259 atcgaatgcgtggtggtgctgtcgtgagacacgttcaactccacgtgaaacgagctgtgac 318
Db 186503 ATCGGATGCCGTGGCGCTCTGCGGGGACAAACGTCACATCCACGTAAGAAAGAGCTGTGAC 186562
Oy 319 ctgcgtggggagactcgttgaccgcgcgcggg 349
|||||
```


JOURNAL

JOURNAL Submitted (11-NOV-2000) **Uta Annenberg Hazen Genome Sequencing Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA**

COMMENT On Feb 3, 2001 this sequence version replaced gi:11138183.

COMMENT On Feb 3, 2001 this sequence version replaced g1:11138183.

Center: Lita Annenberg Hazen Genome Center, Cold Spring Harbor Laboratory

Center code: CSHL

Web site: <http://www.cshl.org/geneseq>

Contact: mccomble@cs.hawaii.edu

Project Information

Contract number: PF33-310P14

Center project name: KP23-310P14
Center clone name: PD23-310P14

Center clone name: KP23-310P14

* NOTE: This is a working draft sequence. It currently consists of 29 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	25187	contlg of 25187 bp in length
2	25188	contlg of 25188 bp in length
3	25189	contlg of 25189 bp in length
4	25190	contlg of 25190 bp in length
5	25191	contlg of 25191 bp in length
6	25192	contlg of 25192 bp in length
7	25193	contlg of 25193 bp in length
8	25194	contlg of 25194 bp in length
9	25195	contlg of 25195 bp in length
10	25196	contlg of 25196 bp in length
11	25197	contlg of 25197 bp in length
12	25198	contlg of 25198 bp in length
13	25199	contlg of 25199 bp in length
14	25200	contlg of 25200 bp in length
15	25201	contlg of 25201 bp in length
16	25202	contlg of 25202 bp in length
17	25203	contlg of 25203 bp in length
18	25204	contlg of 25204 bp in length
19	25205	contlg of 25205 bp in length
20	25206	contlg of 25206 bp in length
21	25207	contlg of 25207 bp in length
22	25208	contlg of 25208 bp in length
23	25209	contlg of 25209 bp in length
24	25210	contlg of 25210 bp in length
25	25211	contlg of 25211 bp in length
26	25212	contlg of 25212 bp in length
27	25213	contlg of 25213 bp in length
28	25214	contlg of 25214 bp in length
29	25215	contlg of 25215 bp in length
30	25216	contlg of 25216 bp in length
31	25217	contlg of 25217 bp in length
32	25218	contlg of 25218 bp in length
33	25219	contlg of 25219 bp in length
34	25220	contlg of 25220 bp in length
35	25221	contlg of 25221 bp in length
36	25222	contlg of 25222 bp in length
37	25223	contlg of 25223 bp in length
38	25224	contlg of 25224 bp in length
39	25225	contlg of 25225 bp in length
40	25226	contlg of 25226 bp in length
41	25227	contlg of 25227 bp in length
42	25228	contlg of 25228 bp in length
43	25229	contlg of 25229 bp in length
44	25230	contlg of 25230 bp in length
45	25231	contlg of 25231 bp in length
46	25232	contlg of 25232 bp in length
47	25233	contlg of 25233 bp in length
48	25234	contlg of 25234 bp in length
49	25235	contlg of 25235 bp in length
50	25236	contlg of 25236 bp in length
51	25237	contlg of 25237 bp in length
52	25238	contlg of 25238 bp in length
53	25239	contlg of 25239 bp in length
54	25240	contlg of 25240 bp in length
55	25241	contlg of 25241 bp in length
56	25242	contlg of 25242 bp in length
57	25243	contlg of 25243 bp in length
58	25244	contlg of 25244 bp in length
59	25245	contlg of 25245 bp in length
60	25246	contlg of 25246 bp in length
61	25247	contlg of 25247 bp in length
62	25248	contlg of 25248 bp in length
63	25249	contlg of 25249 bp in length
64	25250	contlg of 25250 bp in length
65	25251	contlg of 25251 bp in length
66	25252	contlg of 25252 bp in length
67	25253	contlg of 25253 bp in length
68	25254	contlg of 25254 bp in length
69	25255	contlg of 25255 bp in length
70	25256	contlg of 25256 bp in length
71	25257	contlg of 25257 bp in length
72	25258	contlg of 25258 bp in length
73	25259	contlg of 25259 bp in length
74	25260	contlg of 25260 bp in length
75	25261	contlg of 25261 bp in length
76	25262	contlg of 25262 bp in length
77	25263	contlg of 25263 bp in length
78	25264	contlg of 25264 bp in length
79	25265	contlg of 25265 bp in length
80	25266	contlg of 25266 bp in length
81	25267	contlg of 25267 bp in length
82	25268	contlg of 25268 bp in length
83	25269	contlg of 25269 bp in length
84	25270	contlg of 25270 bp in length
85	25271	contlg of 25271 bp in length
86	25272	contlg of 25272 bp in length
87	25273	contlg of 25273 bp in length
88	25274	contlg of 25274 bp in length
89	25275	contlg of 25275 bp in length
90	25276	contlg of 25276 bp in length
91	25277	contlg of 25277 bp in length
92	25278	contlg of 25278 bp in length
93	25279	contlg of 25279 bp in length
94	25280	contlg of 25280 bp in length
95	25281	contlg of 25281 bp in length
96	25282	contlg of 25282 bp in length
97	25283	contlg of 25283 bp in length
98	25284	contlg of 25284 bp in length
99	25285	contlg of 25285 bp in length
100	25286	contlg of 25286 bp in length
101	25287	contlg of 25287 bp in length
102	25288	contlg of 25288 bp in length
103	25289	contlg of 25289 bp in length
104	25290	contlg of 25290 bp in length
105	25291	contlg of 25291 bp in length
106	25292	contlg of 25292 bp in length
107	25293	contlg of 25293 bp in length
108	25294	contlg of 252

FEATURES	
*	202054
*	202285: gap of unknown length
*	202286
*	204454: config of 2169 bp in length
*	204455
*	204686: gap of unknown length
*	204687
*	206723: config of 2037 bp in length
*	206724
*	206953: gap of unknown length
*	206956
*	207411: config of 456 bp in length.
	Location/Qualifiers

[illegible]

Query Match	8.1%	Score 132.4	DB 77	Length 207411
Best Local Similarity	82.1%	Pred. No. 3.5e-15		
Matches 151	Conservative 0	Mismatches 33	Indels 0	Gaps 0

Oy	912	gagaaatctccacagatgctcatctcaacaggaacatcagagagcgccttgcgcccacagc	971
Db	88488	GAGAAATCTCCACAGATGCTTGTGTGTCAATAGAGACTGCACTGCCCTGCGGCAGAGGC	88544
Oy	972	cgctcggagacgagccgcgcagctccacagggcccgcccgccggaaccgcagctcaagccg	1031
Db	88548	CGTTCGGGCTCCACAGCGGGTAACCCGGGGGCCCCGACGCCCAAGAGCCCAAGCTCAGCCG	88607
Oy	1032	ggctcggagcgcctccgaggtccgcgccttgcgcagagacaaagagcctttagctccggagcc	1091
Db	88608	CGCTCTAGAGCGCTCGGTGTGTCCGTGCTGTGCTGACAGACCAAGGGCTTTAGCTCGGGTGA	88667
Oy	1092	gtgg 1095	
Db	88668	GTGG 88671	

RESULT	14
LOCUS	AC087129/c
DEFINITION	AC087129 266574 bp DNA HTG 09-DEC-2000 Mus musculus clone RP23-106A10, WORKING DRAFT SEQUENCE, 30 unordered pieces.
ACCESSION	AC087129
VERSION	AC087129.1 GI:11610855
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT. house mouse.
SOURCE	Mus musculus
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 266574)
AUTHORS	DOE Joint Genome Institute.
TITLE	Sequencing Of Mouse
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 266574)
AUTHORS	DOE Joint Genome Institute.
TITLE	Direct Submission
JOURNAL	Submitted (09-DEC-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT	-----Genome Center


```
* 1 1037: contig of 1037 bp in length
* 1038 1137: gap of unknown length
* 1138 2279: contig of 1142 bp in length
* 2280 2379: gap of unknown length
* 2380 4023: contig of 1644 bp in length
* 4024 4123: gap of unknown length
* 4124 6068: contig of 1945 bp in length
* 6069 6168: gap of unknown length
* 6169 8134: contig of 1966 bp in length
* 8135 8235: gap of unknown length
* 8235 9583: contig of 1349 bp in length
* 9584 9684: gap of unknown length
* 9684 10755: contig of 1072 bp in length
* 10756 10855: gap of unknown length
* 10856 12820: contig of 1965 bp in length
* 12821 12920: gap of unknown length
* 12921 14567: contig of 1647 bp in length
* 14568 14667: gap of unknown length
* 14668 16842: contig of 2175 bp in length
* 16843 16942: gap of unknown length
* 16943 19751: contig of 2809 bp in length
* 19752 19851: gap of unknown length
* 19852 22716: contig of 2865 bp in length
* 22717 22816: gap of unknown length
* 22817 24913: contig of 2097 bp in length
* 24914 25013: gap of unknown length
* 25014 27155: contig of 2142 bp in length
* 27156 27255: gap of unknown length
* 27256 29829: contig of 2574 bp in length
* 29830 29929: gap of unknown length
* 29930 34224: contig of 4295 bp in length
* 34225 34324: gap of unknown length
* 34325 37230: contig of 2906 bp in length
* 37231 37330: gap of unknown length
* 37331 42390: contig of 5060 bp in length
* 42391 42490: gap of unknown length
* 42491 46748: contig of 4258 bp in length
* 46749 46848: gap of unknown length
* 46849 51962: contig of 5114 bp in length
* 51963 52062: gap of unknown length
* 52063 57367: contig of 5305 bp in length
* 57368 57467: gap of unknown length
* 57468 63551: contig of 6084 bp in length
* 63552 63651: gap of unknown length
* 63652 70417: contig of 6766 bp in length
* 70418 70517: gap of unknown length
* 70518 85781: contig of 15264 bp in length
* 85782 85881: gap of unknown length
* 85882 103473: contig of 17592 bp in length
* 103474 103573: gap of unknown length
* 103574 125527: contig of 21954 bp in length
* 125528 125627: gap of unknown length
* 125628 148303: contig of 22576 bp in length
* 148304 148303: gap of unknown length
* 148304 174330: contig of 26027 bp in length
* 174331 174430: gap of unknown length
* 174431 206307: contig of 31877 bp in length
* 206308 206407: gap of unknown length
* 206408 266574: contig of 60167 bp in length.
```

FEATURES

source

```
1..266574
  /organism="Mus musculus"
  /db_xref="taxon:10090"
  /clone_id="RP23-106A10"
```

BASE COUNT 68787 a 61358 c 61225 g 72285 t 2919 others

ORIGIN

Query Match 6.6%; Score 108.2; DB 77; Length 266574;
Best local similarity 77.5%; Pred. No. 7, 1e-11;
Matches 131; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

```
Qy 199 gatctgaacacgctcgaggcgagctcgagcaagtgtcgcgagcgccgacatctgac 258
|||
Db 123551 GATGTGAACACTTGAAGCGCGCCCTGAGTTCGGAAGAGTGTGCAGCACGCCATTGAC 123610
Qy 259 atcgatgcgagcgagctgctcgagcaacgctacatcagctgaagaagactctgac 318
|||
Db 123611 GTTGGAGGCTGTGACTCTGCTCCGCGACAAATGTCACCTCTGTGAAGAAATGAACTATGAC 123670
Qy 319 ctgcgtgaggagctcgtagcccgcccgagctgagagctgtcccgag 367
|||
Db 123671 CTCGCTGGGAGACTCTGTGACCTGCCAGAGGTGAGAGGGGAAACTRAG 123719
```

Search completed: June 28, 2001, 20:34:00
Job time: 6483 sec

Fri Jun 29 08:51:07 2001

us-09-717-883a-1.rge

Page 18

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2001, 19:20:42 ; Search time 146.26 Seconds
(without alignments)
7019.138 Million cell updates/sec

Title: US-09-717-883a-1

Perfect score: 1635
Sequence: 1 tcgcggccgcagagagccgcgc.....agatgctgctccantnaa 1635

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N.Geneseq_0601.*

1: /SIDSI/gcgdata/geneseq/geneeqn/NA1980.DAT:*
2: /SIDSI/gcgdata/geneeq/geneeqn/NA1981.DAT:*
3: /SIDSI/gcgdata/geneeq/geneeqn/NA1982.DAT:*
4: /SIDSI/gcgdata/geneeq/geneeqn/NA1983.DAT:*
5: /SIDSI/gcgdata/geneeq/geneeqn/NA1984.DAT:*
6: /SIDSI/gcgdata/geneeq/geneeqn/NA1985.DAT:*
7: /SIDSI/gcgdata/geneeq/geneeqn/NA1986.DAT:*
8: /SIDSI/gcgdata/geneeq/geneeqn/NA1987.DAT:*
9: /SIDSI/gcgdata/geneeq/geneeqn/NA1988.DAT:*
10: /SIDSI/gcgdata/geneeq/geneeqn/NA1989.DAT:*
11: /SIDSI/gcgdata/geneeq/geneeqn/NA1990.DAT:*
12: /SIDSI/gcgdata/geneeq/geneeqn/NA1991.DAT:*
13: /SIDSI/gcgdata/geneeq/geneeqn/NA1992.DAT:*
14: /SIDSI/gcgdata/geneeq/geneeqn/NA1993.DAT:*
15: /SIDSI/gcgdata/geneeq/geneeqn/NA1994.DAT:*
16: /SIDSI/gcgdata/geneeq/geneeqn/NA1995.DAT:*
17: /SIDSI/gcgdata/geneeq/geneeqn/NA1996.DAT:*
18: /SIDSI/gcgdata/geneeq/geneeqn/NA1997.DAT:*
19: /SIDSI/gcgdata/geneeq/geneeqn/NA1998.DAT:*
20: /SIDSI/gcgdata/geneeq/geneeqn/NA1999.DAT:*
21: /SIDSI/gcgdata/geneeq/geneeqn/NA2000.DAT:*
22: /SIDSI/gcgdata/geneeq/geneeqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	179.6	11.0	202	21 AAC31408	Human secreted pro
2	127.8	7.8	134	16 AAT2548	Human gene signatu
3	93.2	5.7	1710	21 AA295379	Human colon specif
4	93.2	5.7	2220	22 AAF32641	Human CDNA encodin
5	92.2	5.6	3245	14 AAQ40730	Mouse eps8. Mus m
6	53	3.2	114955	20 AAX53491	Human adenosine A1
7	52	3.2	4291	21 AAX75706	Human ORFX ORF1261
8	50.8	3.1	114955	20 AAX53491	Human adenosine A1
9	49.4	3.0	1127	21 AAA02477	Human colon cancer
10	48.2	2.9	1203	20 AA200869	Human C-Maf CDNA.
11	48	2.9	3957	22 AAA09686	HSV-2 Immediate ea

12	47.4	2.9	3957	22 AAA09686.
13	47.2	2.9	1000	21 AAA02484
14	47.2	2.9	1218	21 AAA02488
15	46.6	2.9	7175	16 AA084658
16	46.6	2.9	7175	19 AA042686
17	46.6	2.9	7175	21 AAA71704
18	46.6	2.9	7266	14 AA029059
19	46.6	2.9	7362	19 AA037817
20	46.6	2.9	7362	16 AA084657
21	46.6	2.9	7362	19 AA042685
22	46.6	2.9	7362	21 AA071703
23	46.6	2.9	7376	20 AAX88001
24	46.4	2.8	23187	21 AAA50273
25	46.4	2.8	23187	22 AAF62331
26	46	2.8	8438	15 AA073500
27	46	2.8	24379	18 AAT93095
28	46	2.8	24379	19 AAV25925
29	45.6	2.8	3737	21 AA064656
30	45.6	2.8	3994	21 AAC76475
31	45	2.8	6232	13 AA029269
32	45	2.8	7175	14 AA037818
33	44.6	2.7	2064	14 AA052638
34	44.4	2.7	499	14 AA040729
35	44.4	2.7	1457	22 AAC92642
36	44.4	2.7	2267	20 AAX84665
37	44.4	2.7	10732	21 AAA10594
38	44.2	2.7	22976	20 AAX83426
39	43.8	2.7	489	20 AAX54661
40	43.8	2.7	489	21 AAF20230
41	43.8	2.7	489	21 AAA34108
42	43.8	2.7	6225	20 AAX55273
43	43.8	2.7	6225	21 AAF20843
44	43.8	2.7	6225	21 AAA34721
45	43.8	2.7	21185	21 AAA63350

ALIGNMENTS

RESULT 1	AAC31408	standard: CDNA; 202 BP.
ID	AAC31408	
XX	AAC31408;	
AC		
XX		
DT	06-OCT-2000 (first entry)	
XX		
DE	Human secreted protein 5' EST, SEQ ID NO: 35483.	
XX		
KW	Human; 5' EST; expressed sequence tag; secreted protein; CDNA isolation;	
KM	gene therapy; chromosome mapping; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	EP1033401-A2.	
XX		
PD	06-SEP-2000.	
XX		
PF	21-FEB-2000; 2000BP-0200610.	
XX		
PR	26-FEB-1999; 990S-0122487.	
XX		
PA	(GEST) GENSET.	
XX		
PI	Dumas Mline Edwards J, Duclert A, Giordano J;	
XX		
DR	WPI; 2000-500381/45.	
XX		
PT	New nucleic acid that is a 5' expressed sequence tag (5' EST) for	
PT	obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for	
PT	diagnostic, forensic, gene therapy and chromosome mapping procedures -	
XX		
PS	Claim 1; SEQ ID 35483; 71pp + CD-ROM; English.	

HSV-2 Immediate ea
Human colon cancer
Human colon cancer
Human neuronal cal
DNA encoding human
Human calcium chan
Human calcium chan
Sequence encoding
Human neuronal cal
DNA encoding human
Human calcium chan
N-type calcium cha
Human lipolysis st
Human leptin fragm
DNA encoding pseud
Streptomyces freo
Streptomyces roseo
Partial sequence c
Human ORFX ORF2030
Human calcium chan
Sequence encoding
Streptomyces fradi
Human eps8. Homo
Human NCK-2 CDNA.
Human KDR signal t
Gene encoding a su
Genomic region con
Neutrophil elastase
Human neutrophil e
Human adenosine re
Human enzyme-relat
Human multiple tar
Human adenosine re
Streptomyces globi

XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or poly(A)⁺ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs with intact 5' ends and can therefore be
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 202 BP; 46 A; 44 C; 82 G; 27 T; 3 other;

Query Match 11.0%; Score 179.6; DB 21; Length 202;
Best Local Similarity 96.5%; Pred. No. 1,1e-28;
Matches 191; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

OY 1085 cggagccgtgacgcgcgtggtgtgtctacccgagcgccacatttctcgtcgaagaag 1144
D 4 caggacccgtgacgcgcgtggtgtgtctacccgagcgccacatttctcgtcgaagaag 63
OY 1145 gagctcgagcggtgagcccgagaggggacgctgtgtacagccaggtcacctgcag 1204
D 64 gagctcgagcggtgagcccgagaggggacgctgtgtacagccaggtcacctgcag 123
OY 1205 cgccttcgtctgtgagagcaaaagagtgatcagagctgtgagcagtgatgaggaagca 1264
D 124 cgc-tcgtctgtcgtgagagcaaaagagtgatcagagcgtgagcagtgatgaggaagca 182
OY 1265 aaagaagaagtggaag 1282
D 183 aaagaagaagcggaatg 200

RESULT 2
ID AAT22548
AC AAT22548: standard; cDNA to mRNA; 134 BP.
XX
DT 01-OCT-1996 (first entry)
XX
DE Human gene signature HUMGS04161.
XX
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
XX
OS Homo sapiens.
XX
PN MO9514772-A1.
XX
PD 01-JUN-1995.
XX
PF 11-NOV-1994; 94WO-JP01916.
XX
PR 12-NOV-1993; 93JP-0355504.
XX
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
XX
PI Matsubara K, Okubo K;
XX
DR WPI; 1995-206931/27.
XX
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that

PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
XX
PS Claim 1; Page 1154; 2245pp; Japanese.
XX
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in AAT19001-726837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
XX
SQ Sequence 134 BP; 32 A; 31 C; 47 G; 23 T; 1 other;

Query Match 7.8%; Score 127.8; DB 16; Length 134;
Best Local Similarity 97.7%; Pred. No. 5,7e-18;
Matches 129; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1482 gatctgactgctgctggaggtggtgagggcgctgagacagtcctaggaagagctagcaga 1541
D 1 gatctgactgctgctggaggtggtgagggcgctgagacagtcctaggaagagctagcaga 60
OY 1542 ccccgagaggtgtcaatgtgagccctgagcatgtgtaatatgcgcccagcctataaacag 1601
D 61 ccccgagaggtgtcaatgtgagccctgagcatgtgtaatatgcgcccagcctataaacag 120
OY 1602 cctccgtgctta 1613
D 121 cctccgtgctta 132

RESULT 3
ID AA295379
AC AA295379: standard; DNA; 1710 BP.
XX
DT 01-JUN-2000 (first entry)
XX
DE Human colon specific gene (CSG) SEQ ID NO:1.
XX
KW Human; colon specific gene; CSG; diagnosis; colon cancer; detection;
KW cytostatic; ds.
XX
OS Homo sapiens.
XX
PN WO200007632-A1.
XX
PD 17-FEB-2000.
XX
PF 20-JUL-1999; 99WO-US16357.
XX
PR 04-AUG-1998; 98US-0095231.
XX
PA (DIAD-) DIADEXUS LLC.
XX
PI Sun Y, Recipon H, Macina RA;
XX
DR WPI; 2000-205579/18.
XX
PT Novel methods for diagnosing, monitoring, staging, imaging and treating
PT colon cancer by measuring the level of colon specific gene markers
XX

PS Claim 6; Page 37-38; 42pp; English.

XX The present invention describes a method for diagnosing the presence of
CC colon cancer in a patient. The method comprises measuring levels of
CC colon specific gene markers (CSG) in cells, tissues or bodily fluids,
CC and comparing the measured levels of CSG with levels of CSG from a
CC normal human control, where an increase in measured CSG levels in the
CC patient versus control is associated with the presence of colon cancer.
CC AA295379 to AA295381 represent human CSG sequences. The method is used
CC to detect, monitor, stage or give a prognosis for colon cancer.
CC Antibodies against CSGs are used for detection or image localisation of
CC the CSGs. The antibodies can be conjugated to cytotoxic agent or drug and
CC used to treat colon cancer. The method of the invention is more accurate
CC than prior art clinical methods for staging colon cancer, because it
CC measures colon specific markers, and, unlike pathological staging
CC methods, do not depend on an invasive procedure.

XX
SQ Sequence 1710 BP; 437 A; 529 C; 444 G; 299 T; 1 other:

Query Match 5.7%; Score 93.2; DB 21; Length 1710;
Best Local Similarity 49.1%; Pred. No. 1.3e-10;
Matches 336; Conservative 0; Mismatches 338; Indels 10; Gaps 3;

OY 65 ccctcggagggccagatcacagcagctgctgcagagaatcagtcagccttcagcctgctg 124
DB 431 ctccaccagagcagcagatcagctgcttccttcagagaatcagtcagccttcagcctgctg 490
OY 125 gcccgctgcggcgagacacatcgccagcccttccttcggagctcttgcactctcttc 184
DB 491 ggaagagcttgcagccttgcagagagacagctgcagccttcagtcacacacacacac 550
OY 185 gggcctctgcagatgctgagacagctgcggggcgagctgcagagcagctgctg 244
DB 551 aagtcctcgaattacatcctgcagcagctgcctgcggctgcctgcagcccaagctgac 610
OY 245 cggccgacatctgacatcgatcgctgcgctgcggagacacatccttcacagctgaa 304
DB 611 tcaacccttcacaccccttaagctatcaaccctgcctacagctcgtctacagccacactg 670
OY 305 aacgagctctgagctctgctgcggagacctgcgaccccgccgctgcagctgctcccg 364
DB 671 agtaaccttgcagctgagctgcggccagctgcagcactagccggcgacctgcagacag 730
OY 365 gggagggagggccacacacagcttcttcagcgctgcggagcccgctcacctgac 424
DB 731 gatgagccctctgcctacacacacacatctcagatgactgcgaacttcacagccctcc 790
OY 425 ccgacagagccgc--gctcggagagaccacagctgagaaacagct--acagacagagcg 478
DB 791 agcagaagcacccttagatatacagagccctgttcccttcggcgagagctatagta 850
OY 479 agggcgcgagcagaagcgcccgagctgcgttcacatggtccacccgagagctggagc 538
DB 851 gggagacacacacacatcttcttcagagagacacacacacacacacacacacacac 910
OY 539 cgaagatctgagcctcagctg---gagtcagagacagcaggaatggttcctgtaact 594
DB 911 cccaactcagagccttcacagcccaaacctgcagccgacccctgtaaatatgactctg 970
OY 595 tatgacttcagagcccgcaacagcagctgctgcgttcagacagcgagcagctactgag 654
DB 971 tacgagcttgcagctagagccacagcagcagctgctgctcagagagagagagcagc 1030
OY 655 gccctcggagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 714
DB 1031 gtctcggagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1090
OY 715 gtgcctacacacatcctgcagacc 738
DB 1091 attccagacacacatcctgcagacc 1114

RESULT 4
AAF32641

ID AAF32641 standard; cDNA: 2220 BP.

XX AAF32641;

DT 22-MAR-2001 (first entry)

DE Human cDNA encoding intracellular signalling molecule INTRA4.

XX Human; intracellular signalling molecule; INTRA; immunosuppressive;
KM cytosolic; neuroprotective; neurotropic; antiarteriosclerotic; cancer;
KM antiinflammatory; anti-HIV; neuroleptic; antibacterial; antifungal;
KM antiviral; antiparasitic; antihelminthic; antiparkinsonian; AIDS;
KM cell proliferative disorder; arteriosclerosis autoimmune; epilepsy;
KM inflammatory disorder; Addison's disease; gastrointestinal disorder;
KM neurological disorder; Parkinson's disease; Creutzfeldt-Jakob disease;
KM mental disorder; schizophrenia; anxiety; ss.

XX Homo sapiens.

XX WO200077040-A2.

XX 21-DEC-2000.

XX 16-JUN-2000; 2000WO-US16636.

XX 16-JUN-1999; 99US-0139566.

XX 17-AUG-1999; 99US-0149640.

XX 09-NOV-1999; 99US-0164417.

XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Tang YT, Hillman JL, Lal P, Bandman O, Baughn MR;

XX Azimzal Y, Yang J, Reddy R, Lu DM;

XX WPI: 2001-025334/03.

XX P-PSDB: AAB64372.

XX New human intracellular signalling molecules, useful for the diagnosis,
PT prevention and treatment of cell proliferative, autoimmune,
PT inflammatory, neurological, gastrointestinal, reproductive and
PT developmental disorders -

XX Claim 5; Page 162; 192pp; English.

XX Sequences AAF32638 - AAF32689 represent cDNA encoding human
CC intracellular signalling molecules INTRA1 - INTRA52, represented in
CC AAB64369 - AAB64420. Modulators of the intracellular signalling molecules
CC of the invention exhibit immunosuppressive; cytosolic; neuroprotective;
CC neurotropic; antiarteriosclerotic; antiinflammatory; anti-HIV;
CC neuroleptic; antibacterial; antifungal; antiviral; antiparasitic;
CC antihelminthic; and antiparkinsonian activity. INTRA polypeptides their
CC agonists and antagonists are useful for the treatment of a condition
CC associated with decreased or increased expression of functional INTRA.
CC Disorders associated with abnormal INTRA expression or activity include
CC cell proliferative disorders e.g. arteriosclerosis and cancers;
CC autoimmune or inflammatory disorders e.g. Addison's disease and acquired
CC immunodeficiency syndrome (AIDS); viral, bacterial, fungal, parasitic,
CC protozoal and helminthic infections; gastrointestinal disorders e.g.
CC dysphagia and irritable bowel syndrome; neurological disorders e.g.
CC epilepsy and Parkinson's disease; prion diseases e.g. Creutzfeldt-Jakob
CC disease and mental disorders e.g. anxiety, schizophrenia and Tourette's
CC disorder. Antibodies immuno specific for the INTRA proteins may also be
CC useful in the diagnosis of the above disorders.

XX Sequence 2220 BP; 549 A; 696 C; 593 G; 382 T; 0 other:

Query Match 5.7%; Score 93.2; DB 22; Length 2220;
Best Local Similarity 49.1%; Pred. No. 1.3e-10;
Matches 336; Conservative 0; Mismatches 338; Indels 10; Gaps 3;

```

OY 65 cccctggagagccgagttacacgagctgtcgacgaagatcaagtaacgcttcagctgtctg 124
DB 948 ctcaccacccagccacagttatgtactgtcttcagaaagatcaagtaacgcttcagctgt 1007
OY 125 gcccgctgtcgaggaacacatcgccaccctctctccggaactgtttgactctcttc 184
DB 1008 ggaaggtcggcactctgtgtgaagagaagatgcccttgatctcgttaacatctcttc 1067
OY 185 gggcctctgacatgtatgttgacacgctcgggggggccggaattcgcgagcagtgcg 244
DB 1068 aagtcctgaattcatctcgtcgaggtgcctcgaggctgctagcagcccaagtgatc 1127
OY 245 cggcgcatctgacatcgatcgatgcgtgctgctgctgctggggaacgtcaactcaag 304
DB 1128 tcacccctctaccccttaagaactacacactgctacagctctgtcttaagccactgag 1187
OY 305 aacgagctctgacatctgctgtggggagctgtgaccccgcccggtctgagctgtcccg 364
DB 1188 agtaaccttggatgtgggttggccacgctggaaccactagccggccgacatggaagc 1247
OY 365 gaggaggaaccccaacagacccgagttctcaagcggctgggaagcggcgctacgtac 424
DB 1248 gatgagccctctgctacacacacacacatctcgtgagtactggaactccagagccctcc 1307
OY 425 ccgagagagcgc---gcttggagagacccagttgagaacagct---acagcagagcgg 478
DB 1308 agccaagacaccttagatgatacagagacccctgttcccttcggggggaagtatagttta 1367
OY 479 aggcgcggcgcaagaagcgcccccgaagtcgctgcaatgttcacggagagacttgaggc 538
DB 1368 gggagacactcaacttctctcagagaagacacacacacacacacacacacacacacac 1427
OY 539 cagaatctgagcctagctgctgctgctgctgctgctgctgctgctgctgctgctgctg 594
DB 1428 cccaactcagagccctcagagccacacacacacacacacacacacacacacacacacac 1487
OY 595 tatgcttccggagcccgacagacagcagtgatgctgctgctgctgctgctgctgctgctg 654
DB 1488 taagagtttgaagctagagaaacacagggaaactgactgtgtccagggagagaaagctgag 1547
OY 655 gtccctgagtagcagtcgtgaagtgctggaaggttcgggacccagcgggcgaggaagga 714
DB 1548 gtctggagccacagcaagcggtggtgctgtggaagatgagggcgagcgagcgctac 1607
OY 715 gtgcctcacaatctctacacc 738
DB 1608 attccaagcaacatctctgagccc 1631

```

RESULT 5
AA040730
ID AA040730 standard; DNA: 3245 BP.

```

AC AA040730;
XX
XX 25-AUG-1993 (first entry)
XX
XX Mouse eps8.
XX
XX Epidermal growth factor receptor; EGFR-pathway substrate; eps;
XX tyrosine kinase receptor; TKR; SH2; SH3; mitogenesis; ss.
XX
XX Mus musculus.
XX
XX Key Location/Qualifiers
XX CDS 246..2711
XX FT /*tag= a
XX FT polyA_signal 3036..3041
XX FT /*tag= b
XX
XX US7935311-A.
XX
XX PD 01-APR-1993.

```

```

XX
XX 25-AUG-1992; 92US-0935311.
XX
XX 25-AUG-1992; 92US-0935311.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX
XX DI Fiore PP, Fazioli F;
XX
XX WPI: 1993-159477/19.
XX
XX P-PSDB: AAR35451.
XX
XX Epidermal growth factor receptor substrate, eps 8 - used to
XX enhance mitogenic response of cells to epidermal growth factor
XX
XX Disclosure; Page 30-34; 40pp; English.
XX
XX Eps8 is a novel EGFR substrate. The protein bears the
XX characteristic signatures of TKR substrates including SH2 and
XX SH3 domains. Eps8 is involved in the transduction of mitogenic
XX signals and it can be used to enhance the mitogenic response of
XX cells to EGF.
XX
XX Sequence 3245 BP; 885 A; 805 C; 865 G; 690 T; 0 other;

```

Query Match 5.6%; Score 92.2; DB 14; Length 3245;
Best Local Similarity 51.9%; Pred. No. 2.2e-10;
Matches 233; Conservative 0; Mismatches 213; Indels 3; Gaps 1;

```

OY 30 ctggggagggctctgtcagctcgtcgagcgaagccgctcgtgagggcgaatcacgag 89
DB 1177 ctggagagggggtttaaactagagggcaaacgccacccctcgcagagtttgact 1236
OY 90 tgcgaggaagaataagttacgctcagcgtgctgctggccggctgcggaacgtgcg 149
DB 1237 gtccacagaatttaaacatgattcaactcttcggccaagtgaatccacataccaga 1296
OY 150 acccctctcctcgagagctgtgtcactctcttcgggctctgcaatgattgagaca 209
DB 1297 accgagtgctcagatcgtgttcatctttgttactcactaataatgtgtccagg 1356
OY 210 cgtcgggggggcggaattcgcgagcagtgctgcgcgccgcatctacatcgatgcg 269
DB 1357 caaaggttggccctgaactggtccagttcgttaccagccactgtgacaagaacagacag 1416
OY 270 tggcgctgctcgaggaacgctacacagctggaacgagctgtgactcgtgtggg 329
DB 1417 ttgatttcttaactacacagccactgcggaagagaagcgtgtgagatgctcagtgag 1476
OY 330 actcgtggaacccgcccggctggaagctgtcccccggagag---gaaacccacatacagac 386
DB 1477 atagttgtgtaaggtgagagcagagtggtcggaagagacagttacatccacttaagtc 1536
OY 387 ccgagttctcagcgcgtctggaagcgcggtacactgaccccgagagcgcgctgtgagag 446
DB 1537 cgaagttccgaaagcgtgtgggaccccgatgtcgtgaacttcattggtggcgccacagagc 1596
OY 447 acccagttggaacacagctacagcagag 475
DB 1597 aagacatgtatcaacttgcgcgagttcgtg 1625

```

RESULT 6
AA053491/C
ID AA053491 standard; DNA: 114955 BP.

```

AC AA053491;
XX
XX 05-JUL-1999 (first entry)
XX
XX Human adenosine A1 receptor antisense oligonucleotide fragment.
XX
XX

```


Dd	543	CCNCCNCCTCCCNCCNCCCNCNCCNCCNCCGCGCNCNCCNCCCGAGGC	484
Oy	959	cctgcccagaggccgcgtggagcagcgacgtccagggccccgcgccgaacc	1018
Dd	483	CTNNCCNCCNCCGNCCCCCNCGCTTCNCCNCCNCCNCCNCCNCCNCCNCC	424
Oy	1019	gcagctcaccccgagtcgcgacgctcc	1046
Dd	423	CCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC	396
 RESULT 14 AAA02488/c ID AAA02488 standard; cDNA: 1218 BP. XX AC AA02488; XX DT 19-MAY-2000 (first entry) Dx Human colon cancer cell line polynucleotide sequence SEQ ID NO:2479. XW Human; colon cancer; tumour; diagnosis; gene expression product; KW probe; detection; cancerous state; metastasis; identification; KW breast cancer; oestrogen receptor-positive breast cancer; therapy; KW oestrogen receptor-negative breast cancer; lung cancer; ss. XA Homo sapiens. OS XX MO958675-A2. PN PD 18-NOV-1999. XX FE 13-MAY-1999; 99WO-US10602. XX XX 14-MAY-1998; 98US-0085426. PR 15-MAY-1998; 98US-0085537. PR 15-MAY-1998; 98US-0085696. PR 21-OCT-1998; 98US-0105234. PR 27-OCT-1998; 98US-0105877. XA PA (CHIR) CHIRON CORP. PA (HYSE-) HYSEQ INC. XX PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J; PI Reinhard C, Glase K, Randazzo F, Kennedy GC, Pol D, Kassam A; PI Lanson G, Drmanac R, Kravtsov R, Dickson M, Drmanac S, Labat I; PI Leschowitz D, Kita D, Garcia V, Jones LM, Stache-Crain B; XX DR WPI: 2000-126369/11. PT PT Polynucleotide library used to determine cancerous states of mammalian PS cells - PS Claim 1; Page 995-996; 1097pp; English.			
 AAA0010 to AAA02716 represent polynucleotides isolated from cDNA libraries constructed from human colon cancer cell lines. The present invention also describes a method of detecting differentially expressed genes correlated with a cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test sample derived from a cell suspected of being cancerous, where detection of the differentially expressed gene product is correlated with a cancerous state of the cell from which the test sample was derived. The polynucleotide sequences can be used in a method for detecting differentially expressed genes correlated with a cancerous state of a mammalian cell. The polynucleotides can also be used as probes for detecting and mapping related genes. They can be used in diagnosis and prognosis of diseases and disorders (e.g. identification of pre-metastatic or metastatic cancerous states, stages of cancer, or responsiveness of cancer to therapy). This is particularly for breast cancer, oestrogen receptor-positive breast cancer, oestrogen receptor- negative breast cancer, lung cancer, and colon cancer.			

xx	Sequence	1218	BP; 9	A; 31	C; 494	G; 37	T; 647	other;	
xx	Query Match	2.98;	Score	47.2;	DB	21;	Length	1218;	
xx	Best Local Similarity	27.1%;	Pred. No.	0.4;					
xx	Matches	95;	Conservative	0;	Mismatches	256;	Indels	0;	Gaps
Qy	720	ctacaac	ctctg	caccctc	accctc	ggagcc	cggtc	gcacac	caagc
Db	1081	CNCNNNCNNNN	CNNNCNNNCNNNC	CNNNCNNNCNNNC	CNNNCNNNNNNNNNNNN	CCCCC	10222		
Qy	780	cagcct	taacag	acac	ctc	ctccac	ccac	gcccc	cgccg
Db	1021	CCCCCCCCNNNN	CNNNCNNNCNNNC	CCCCCCCC	CCCCCCCC	CNNNCNNNCNNNNNNNN	NNNN	962	
Qy	840	ggtc	cgcc	cccc	cgctc	gtgga	acag	cccc	gtgga
Db	961	NNCCCCCCCC	CCCCC	CNNNC	CCCCC	CNNNNNNNNNN	NNNN	CCCCCCCC	CNNNNNNNC
Qy	900	cccc	agc	gaga	gga	aat	ctcc	acat	gc
Db	901	CNNNC	CCCCC	CNNNN	CNNNN	CNNNN	CNNNN	CCCCC	CNNNNNNNN
Qy	960	ctgc	gccc	acag	gcgc	gtc	ggag	ccg	gcgc
Db	841	CCCCC	CNNNC	CCCCC	CCCCC	CNNNC	CCCCC	CNNNN	CCCCC
Qy	1020	cagct	cac	gcgc	cccc	ggc	gtc	gcgc	gtc
Db	781	NCNNNC	CNNNC	CNNNN	CNNNN	CCCCC	CCCCC	CNNNC	CCCCC
RESULT: 15									
xx	AA084658								
xx	AA084658	standard;	DNA;	7175	BP.				
xx	01-DEC-1995	(first entry)							
xx	Human neuronal calcium channel subunit alpha 1B-2.								
xx	Calcium channel subunit; antagonist; agonist; diagnosis;								
xx	Lambert Eaton Syndrome; ss.								
xx	Homo sapiens.								
xx	OS								
xx	Key	Location/Qualifiers							
xx	CDS	144..6857							
xx	misc_feature	6633..7175							
xx	FT	/*tag= b							
xx	FT	/note= "identical to alpha 1B-1"							
xx	MO9504822-A.								
xx	16-FEB-1995.								
xx	11-AUG-1994;	94MO-US09230.							
xx	11-AUG-1993;	93US-010536.							
xx	05-NOV-1993;	93US-0149097.							
xx	(SALK) SALK INST BIOTECHNOLOGY IND ASSOC.								
xx	Ellis SB, Gillespie A, Harpold MW, McCue AF, Williams ME;								
xx	WPI: 1995-090900/12.								
xx	P-PSDB; AAR71006.								
xx	DNA encoding human calcium channel sub-unit(s) - used for								
xx	developing prods. for studying calcium channels, e.g. for								

PT obtaining agonists and antagonists

XX

PS Disclosure: Page 149-160; 285pp; English.

XX

CC DNA encoding the alpha 1B subunit was isolated by screening a
CC human basal ganglia cDNA library with fragments of the rabbit
CC skeletal muscle calcium channel alpha 1 subunit-encoding cDNA.
CC A portion of one of the positive clones was used to screen an
CC IMR32 cell cDNA library. Clones that hybridised to the basal
CC ganglia probe were used to further screen an IMR32 cell cDNA
CC library to identify overlapping clones that in turn were used
CC to screen a human hippocampus cDNA library. A series of clones
CC to span nearly the entire length of the nt. sequence encoding
CC the human alpha 1B subunit was obtained. Nucleic acid amplification
CC of specific regions of the IMR32 cell alpha 1B mRNA yielded
CC additional segments of the alpha 1B coding sequence. A full-
CC length alpha 1B cDNA clone was constructed by ligating portions
CC of the partial cDNA clones. Nucleic acid amplification analysis
CC of IMR32 cell RNA and genomic DNA using oligo primers corresp. to
CC sequences located 5' and 3' of the stop codon of the DNA encoding
CC the alpha 1B subunit revealed an alternatively spliced alpha
CC 1B-encoding mRNA in IMR32 cells. This second mRNA product is the
CC result of differential splicing of the alpha 1B subunit transcript
CC to include another exon that is not present in the mRNA corresp.
CC to the other 3' alpha 1B cDNA sequence that was initially isolated.
CC The alpha 1B subunit encoded by a DNA sequence contg. an additional
CC exon is referred to as alpha 1B-1 and given in AA084657/R71005.
CC Whereas the other form is referred to as alpha 1B-2 and is given in
CC AA084658/R71006. Following the sequence of the additional exon in
CC alpha 1B-1 the alpha 1B-1 and alpha 1B-2 sequences are identical.
XX

S0 Sequence 7175 BP; 1415 A; 2197 C; 2168 G; 1395 T; 0 other;

Query Match 2.9%; Score 46.6; DB 16; Length 7175;

Best Local Similarity 46.7%; Pred. NO. 0.71;

Matches 209; Conservative 0; Mismatches 237; Indels 2; Gaps 2;

OY 855 ggaacgggcccgcgtggaacgtgtgatatnctcaacggcttgaccacgagagaag 914
DB 2748 gcaaggggcccgaagagcggaagcgggagccggctcccggaagagcgccg 2807
OY 915 aaattctccagatgcatcgtcaacgaagaaactgcagcgcgccctggcccaaggccgc 974
DB 2808 caccgcagcacaagaagaagcgcgcgcccgcccgagcgcgagagcgccga 2867
OY 975 tcgggaacgagcgcgacgtcccaaggcccgcccgcccgaaacgagctcagccgggc 1034
DB 2868 ggcacagggcccgaaggcgcgccgacacacggcgcgctcccgagagggcgcc 2927
OY 1035 tcggacgcttcgaaagtctcgagctgtgacagncaaaggctttagctcggagcgtg 1094
DB 2928 gagcgggagcccgacgcacacggcgcgacgcagcagatccgacaaagagtgccgc 2987
OY 1095 gacgcgcgggtgtgtacacggcgggcgaccttctcgtgcgaagaaggagagctgcg-9 1153
DB 2988 ggcgcgaaggcgagcgcgcgcgcgaccccgcgcccgcccgagcgccggag 3047
OY 1154 gcggtgagaccccgagaggggcaagctgtgtacagcgaaggtcaacggtcagcgcttcgct 1213
DB 3048 gcggaagagcggaagagccggcgcgccgcgcgcggccgacaaagcgcgagcctgtc 3107
OY 1214 gctggaagacaagaaggtctcagagcttgagagcagtgatgtagaagcaaaagaagaa 1273
DB 3108 -caggaagctgtggaagaaggaacacgagagagaggtccacggaagaagaggtcagat 3166
OY 1274 ggtggaagcgaggtggaatggaagtc 1301
DB 3167 agtggaaagccgacgaagaagagagctc 3194

Job time: 4112 sec

...


```

OY 30 ctgggagaggtctgtcagcgtcgccgagccgcccctcggagccgagtaacacgag 89
DB 1177 CTGGAGAGGCGCTTTAAACATGAGGCAAAACCGCACCTCTCTGACGATTTGTGACT 1236
OY 90 tgcgcagaagaatcaagtaacgtcctcagcctgtgtgcccgcgtcgcgcaacatcccg 149
DB 1237 GTTCCAGAAAGTTTAAACATGATTCACCTCTGCGCAAGTTGAAGTCCATATCCAGA 1296
OY 150 accctctctcgcagagctgttgcacctcttccggcctctgcagatgattgtgaaca 209
DB 1297 ACCCGAGTCTCTAGATCTGATTCATTTTGTTCATCTCACTAATAATATGAGTCCAGG 1356
OY 210 cgtcgaggggcgagcagttcgcagcagtgctgcggcgccgcatctgagatcgatgcgcg 269
DB 1357 CAACAGAGTGGCCTTGAACAGTGGCCAGTCTGCTACTCAGCCCACTGTTGACAAAAGCACAG 1416
OY 270 tggcgtctgtcgaggacaacagtcactccacgtgaacacagcctctgcagctcgtcgagg 329
DB 1417 TTGATTTCTTAAACATACACAGCCACTGCGGAGGAACGAACTGTGATGATCTACGTGAG 1476
OY 330 actcgtgacccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 386
DB 1477 ATAGTTGGGTGAAGTGAAGAGCAGAGTGGCCGAAAGAACAGTTCAATCCACCTTACGTC 1536
OY 387 ccgagttctcagcgcgtcgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 446
DB 1537 CGAGGTTCCGCAACGCGCTGGAGCCCCCGATCTGAATCTCATGAGGCGCGCCACACAGAC 1596
OY 447 acccagltgagaacagctacagcagag 475
DB 1597 AAGACATGTATCACTGGCGGAGTCCGCG 1625

```

```

RESULT 2
US-08-368-079-3
; Sequence 3, Application US/08368079
; Patent No. 5610018
; GENERAL INFORMATION:
; APPLICANT: DI Fiore, Pier Paolo
; APPLICANT: Fazio, Francesco
; TITLE OF INVENTION: eps8, A Substrate for the Epidermal Growth Factor
; TITLE OF INVENTION: Kinase, Antibodies Thereof, and Methods of Use Thereof
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/368,079
; FILING DATE: 03-JAN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/935,311
; FILING DATE: 25-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NH035,001DVL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3245 base pairs
; TYPE: nucleic acid

```

```

; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 246..2708
; US-08-368-079-3

```

```

Query Match 5.6%; Score 92.2; DB 1; Length 3245;
Best Local Similarity 51.9%; Pred. No. 1.9e-11;
Matches 233; Conservative 0; Mismatches 213; Indels 3; Gaps 1;

```

```

OY 30 ctgggagaggtctgtcagcgtcgccgagccgcccctcggagccgagtaacacgag 89
DB 1177 CTGGAGAGGCGCTTTAAACATGAGGCAAAACCGCACCTCTCTGACGATTTGTGACT 1236
OY 90 tgcgcagaagaatcaagtaacgtcctcagcctgtgtgcccgcgtcgcgcaacatcccg 149
DB 1237 GTTCCAGAAAGTTTAAACATGATTCACCTCTGCGCAAGTTGAAGTCCATATCCAGA 1296
OY 150 accctctctcgcagagctgttgcacctcttccggcctctgcagatgattgtgaaca 209
DB 1297 ACCCGAGTCTCTAGATCTGATTCATTTTGTTCATCTCACTAATAATATGAGTCCAGG 1356
OY 210 cgtcgaggggcgagcagttcgcagcagtgctgcggcgccgcatctgagatcgatgcgcg 269
DB 1357 CAACAGAGTGGCCTTGAACAGTGGCCAGTCTGCTACTCAGCCCACTGTTGACAAAAGCACAG 1416
OY 270 tggcgtctgtcgaggacaacagtcactccacgtgaacacagcctctgcagctcgtcgagg 329
DB 1417 TTGATTTCTTAAACATACACAGCCACTGCGGAGGAACGAACTGTGATGATCTACGTGAG 1476
OY 330 actcgtgacccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 386
DB 1477 ATAGTTGGGTGAAGTGAAGAGCAGAGTGGCCGAAAGAACAGTTCAATCCACCTTACGTC 1536
OY 387 ccgagttctcagcgcgtcgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 446
DB 1537 CGAGGTTCCGCAACGCGCTGGAGCCCCCGATCTGAATCTCATGAGGCGCGCCACACAGAC 1596
OY 447 acccagltgagaacagctacagcagag 475
DB 1597 AAGACATGTATCACTGGCGGAGTCCGCG 1625

```

```

RESULT 3
PCT-US93-07996-3
; Sequence 3, Application PC/TUS9307996
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States, as represented by the
; APPLICANT: Secretary of Health and Human Services
; TITLE OF INVENTION: Substrate of the Epidermal Growth Factor Kinase
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07996
; FILING DATE: 19930825
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 3:

```



```

1 MEDIUM TYPE: Diskette
2 COMPUTER: IBM Compatible
3 OPERATING SYSTEM: DOS
4 SOFTWARE: FASTSEQ Version 1.5
5 CURRENT APPLICATION DATA:
6 APPLICATION NUMBER: US/08/455.543A
7 FILING DATE: May 31, 1995
8 PRIOR APPLICATION DATA:
9 APPLICATION NUMBER: 08/223,305
10 FILING DATE: April 4, 1994
11 PRIOR APPLICATION DATA:
12 APPLICATION NUMBER: 07/868,354
13 FILING DATE: April 10, 1992
14 PRIOR APPLICATION DATA:
15 APPLICATION NUMBER: US 07/745,206
16 FILING DATE: 15-AUG-1991
17 PRIOR APPLICATION DATA:
18 APPLICATION NUMBER: US 07/620,250
19 FILING DATE: 30-NOV-1990
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER: US 07/482,384
22 FILING DATE: 20-FEB-1990
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: US 07/603,751
25 FILING DATE: 04-APR-1989
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER: WO PCT/US89/01408
28 FILING DATE: 04-APR-1989
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: US 07/176,899
31 FILING DATE: 04-APR-1988
32 ATTORNEY/AGENT INFORMATION:
33 NAME: Seidman, Stephanie L.
34 REGISTRATION NUMBER: 33,779
35 REFERENCE/DOCKET NUMBER: 632-52517
36 TELECOMMUNICATION INFORMATION:
37 TELEPHONE: (619)238-0999
38 TELEFAX: (619)238-0062
39 INFORMATION FOR SEQ ID NO: 8:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 7175 base pairs
42 TYPE: nucleic acid
43 STRANDEDNESS: double
44 TOPOLOGY: linear
45 MOLECULE TYPE: DNA (genomic)
46 FEATURE:
47 NAME/KEY: CDS
48 LOCATION: 144..6857
49 FEATURE:
50 NAME/KEY: 5'UTR
51 LOCATION: 1..143
52 FEATURE:
53 NAME/KEY: 3'UTR
54 LOCATION: 6855..7175
55 US-08-455-543A-8

```

Accession	Gene	Protein	Length
OY 1035	tcgagagcctccgagatccagcgctctgctcgaagcaagagccttagctcccgagacgtg	1094	
Db 2928	GAGGGGAGCCCCCAGCCCAACCCGGGCAACCGGCACCAAGATCCGACAAAGATGCCCC	2987	
OY 1095	gaacgcgtggtgtgtgtctgtacccgagcgcaactttctcgtcagaagaagagatctcg-g	1153	
Db 2988	GAGCCCAAGGGCCGAGCCGGCCGGCCGCGACACCGCGGGGCCCGAGCGGGGCCCGGAG	3047	
OY 1154	gcggtgagcccccagagagagggagcgtgtgtacagccaggtcacccgttcagcgtctgct	1213	
Db 3048	GCGGAGAGCGGGGAGAGGAGCGCGCGCCGGCAGACGGGCCCCGACAAAGCCAGCCTCT	3107	
OY 1214	gctgagaggaacaagaagaagtgtctagagctgtgagcgatgtatgsgaagcaaaagaaga	1273	
Db 3108	-CAcGAGGCTGTGTGAGAGAGAGACCAcCGGAGAAAGcAGGCCAGGAGAAAGAGcTcTcT	3166	
OY 1274	ggtgaaagcgccaggtgtaaatgtgaagttc	1301	
Db 3167	AGTGAAGCCGACAAAGAAAAGAGCTTC	3194	

```

RESULT      6
US-08-193-078B-8
: Sequence 8, Application US/08193078B
: Patent No. 5846757
:
: GENERAL INFORMATION:
: APPLICANT: Harpold, Michael
: APPLICANT: Ellis, Steven
: APPLICANT: Williams, Mark
: APPLICANT: Feldman, Daniel
: APPLICANT: Mocce, Ann
: APPLICANT: Brenner, Robert
: TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
: TITLE OF INVENTION: METHODS
: NUMBER OF SEQUENCES: 29
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BROWN, MARTIN, HALLER & MCCLAIN
: STREET: 1660 UNION STREET
: CITY: SAN DIEGO
: STATE: CA
: COUNTRY: USA
: ZIP: 92101
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/193,078B
: FILING DATE: 07-FEB-1994
: CLASSIFICATION: 435
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/868,354
: FILING DATE: 10-APR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/745,206
: FILING DATE: 15-AUG-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Seidman, Stephanie L.
: REGISTRATION NUMBER: 33,779
: REFERENCE/DOCKET NUMBER: 6362-53607
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-238-0999
: TELEFAX: 619-238-0062
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7175 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS

```

```

; LOCATION: 144..6857
;
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..143
;
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 6855..7175
;
US-08-193-07AB-8

```

Query Match	2.98;	Score	46.6;	DB	2;	Length	7175;
Best Local Similarity	46.78;	Pred.	No.	0.16;			
Matches	209;	Conservative	0;	Mismatches	237;	Indels	2;
						Gaps	2

QY	855	ggaacagccccgccttggagacacgcgcgcataaccacacgccttggaccccgagcgaagaagag	914
Db	2748	GCACGCGCCCCGAAGCGCGAGACCGGGGAGCCCGGTCCCGCGAGACGCCGCCCGCGCGC	2807
QY	915	aaattctccagatgctcatcgttcaacagagaacttgcagcgcgccttgcgcgcgaagggccgc	974
Db	2808	CACCGCCACCCACAGCAAGAGCGCCCGCGGGCCCCCGCGAGCCGAGACCGCGCGCGGA	2867
QY	975	tcggagaccgagcgcgcgagtcctcccaaggcgccgcgcgcgcgcgaacccgagctcaagccgggc	1034
Db	2868	GGCCCAAGGCCCCAGAGCGCGCCGCGCCGCAACACCGCGCGCTCCCCGAGAGAGCGGCC	2927
QY	1035	tcggagcccttcgagnttccgcgcgccttgcgcgcagagacaaagcgtttacgttccggagacgtg	1094
Db	2928	GAGCGGAGGCCCCGAGCGCCACCCCGCGCAACCGCAACAGATCCGAGCAAGAGATGGCGCC	2987
QY	1095	gaagcgcctgggtgtgtctgtacccggcgccacttcttcgtctgcagaaagagaagatctgcg-g	1153
Db	2988	GGCCCCAAGGGCCACACCGCGCCCGCGGCAACCGCGCGCCCCCGAGCGGCGCCCGGAG	3047
QY	1154	ggcgctgagcccgagagagggcgacgtgtgtacagccaggtcacccgttcagcgcgttcgct	1213
Db	3048	GGCGAGAGCGGGGAGAGCGCGCGCGCGGCGCACCGGCGCCGGCCACAAAGCCGACTTCT	3107
QY	1214	gcttgagagacaaagagaaagtgtctcagagcttggagggcagtgatctgagaaagcaaaagaagaa	1273
Db	3108	-CACGAGCGCTGTGGAGAGAGACACACGAGAAAGAGCGCCACGAGAAAGAGAGCTGTACAT	3166
QY	1274	ggctgaaagcgaggtcgaaatctgaatc	1301
Db	3167	AGTGGAGCCGACAGGAAAAAGAGCTTC	3194

```

1  RESULT 7
2  US-08-223-305C-8
3  : Sequence 8, Application US/08223305C
4  : Patent No.5851824
5  :
6  : GENERAL INFORMATION:
7  :
8  : APPLICANT: Harpold, Michael
9  : APPLICANT: Ellis, Steven
10 : APPLICANT: Williams, Mark
11 : APPLICANT: Feldman, Daniel
12 : APPLICANT: Mocue, Ann
13 :
14 : TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
15 : METHODS
16 :
17 : NUMBER OF SEQUENCES: 57
18 :
19 : CORRESPONDENCE ADDRESS:
20 : ADDRESSEE: Brown, Martin, Haller & Mcclain
21 : STREET: 1660 Union Street
22 : CITY: San Diego
23 : STATE: California
24 :
25 : COUNTRY: USA
26 :
27 : ZIP: 92101-2926
28 :
29 : COMPUTER READABLE FORM:
30 : MEDIUM TYPE: Diskette
31 : COMPUTER: IBM Compatible
32 : OPERATING SYSTEM: DOS
33 : SOFTWARE: FASTSEQ Version 1.5
34 :
35 :

```

```

1 CURRENTAPPLICATION DATA:
2 APPLICATION NUMBER: US/08/223.305C
3 FILING DATE: April 4, 1994
4 PRIOR APPLICATION DATA:
5 APPLICATION NUMBER: 07/868,354
6 FILING DATE: April 10, 1992
7 PRIOR APPLICATION DATA:
8 APPLICATION NUMBER: US 07/745,206
9 FILING DATE: 15-AUG-1991
10 PRIOR APPLICATION DATA:
11 APPLICATION NUMBER: US 07/620,250
12 FILING DATE: 30-NOV-1990
13 PRIOR APPLICATION DATA:
14 APPLICATION NUMBER: US 07/482,384
15 FILING DATE: 20-FEB-1990
16 PRIOR APPLICATION DATA:
17 APPLICATION NUMBER: US 07/603,751
18 FILING DATE: 04-APR-1989
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER: WO PCT/US89/01408
21 FILING DATE: 04-APR-1989
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER: US 07/176,899
24 FILING DATE: 04-APR-1988
25 ATTORNEY/AGENT INFORMATION:
26 NAME: Seidman, Stephanie L.
27 REGISTRATION NUMBER: 33,779
28 REFERENCE/DOCKET NUMBER: 52516 (P519739)
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: (619)238-0999
31 TELEFAX: (619)238-0062
32 INFORMATION FOR SEQ ID NO: 8:
33 SEQUENCE CHARACTERISTICS:
34 LENGTH: 7175 base pairs
35 TYPE: nucleic acid
36 STRANDEDNESS: double
37 TOPOLOGY: linear
38 MOLECULE TYPE: DNA (genomic)
39 FEATURE:
40 NAME/KEY: CDS
41 LOCATION: 144..6857
42 FEATURE:
43 NAME/KEY: 5'UTR
44 LOCATION: 1..143
45 FEATURE:
46 NAME/KEY: 3'UTR
47 LOCATION: 6855..7175
48
49 US-08-223-305C-8

```

Query Match	Similarity	2.9%	Score 46.6	DB 2	Length 7175
Best Local	Similarity 46.7%	Pred. No. 0.16			
Matches 209	Conservative 0	Mismatches 237	Indels 2	Gaps 2	
QY	855 gacagcgcgcgctgaggacagctgcgatanctcaacgctctgagccacgagaaagag	914			
Db	2748 GCAGAGAGCCCCCAAGACCGAGACCGGGAGCCCGGTCGCCGGAGAGAGCGCCGCGCGG	2807			
QY	915 aaatctccagatgtctatcgctcaacgaggaactgcagcgcgctctgcccagagcgcg	974			
Db	2808 CACCCGACGACACAGCAAGAGGCGCGGGGGCCCCCGAGAGCGCGAGCCAGCCGGCCGA	2867			
QY	975 tcggagaccgagccgcgcagctcccaaggcccgccgcgcgcggaacgcgagctcaagccggcg	1034			
Db	2868 GCCCCAGGCCCCGAGGCGCGCGCGCCGACACACCGCGCGCGCTCCCGGAGAGCGCGCC	2927			
QY	1035 tcggaagcctcccgangtcgcgcctgcgctgcgagngnaagagcctttagctccggagacgctg	1094			
Db	2928 GAGCGGAGACCCCGAAGCGCCACCGCGCGCACCGCACACAGATCCGACGAAGAGTCCGCC	2987			
QY	1095 gaagcgctggtgtgtctgaacgcgggagcgaactttctcgtctcgaaagaggaagctcg_g	1153			
Db	2988 GCGCCCAAGAGGAGAGGCGCGCGCGCCGACCGCGCGCGCCCCGACCGGGCCCCCGGAG	3047			


```

1 APPLICATION NUMBER: 08/105,536
2 FILING DATE: 11-AUG-1993
3 PRIOR APPLICATION DATA:
4 APPLICATION NUMBER: 07/914,231
5 FILING DATE: 13-JULY-1992
6 PRIOR APPLICATION DATA:
7 APPLICATION NUMBER: 07/868,354
8 FILING DATE: 10-APR-1992
9 PRIOR APPLICATION DATA:
10 APPLICATION NUMBER: PCT/US92/06903
11 FILING DATE: 14-AUG-1992
12 PRIOR APPLICATION DATA:
13 APPLICATION NUMBER: 07/745,206
14 FILING DATE: 15-AUG-1991
15 PRIOR APPLICATION DATA:
16 APPLICATION NUMBER: 07/620,250
17 FILING DATE: 30-NOV-1990
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER: 07/603,751
20 FILING DATE: 08-NOV-1990
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: 07/482,384
23 FILING DATE: 02-FEB-1990
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: PCT/US89/01408
26 FILING DATE: 04-APR-1989
27 PRIOR APPLICATION DATA:
28 APPLICATION NUMBER: 07/176,899
29 FILING DATE: 04-APR-1988
30 ATTORNEY/AGENCY INFORMATION:
31 NAME: Seidman, Stephanie L.
32 REGISTRATION NUMBER: 33,779
33 REFERENCE/DOCKET NUMBER: 6362-519812
34 TELECOMMUNICATION INFORMATION:
35 TELEPHONE: (619) 238-0999
36 TELEFAX: (619) 238-0062
37 INFORMATION FOR SEQ ID NO: 8:
38 SEQUENCE CHARACTERISTICS:
39 LENGTH: 7175 base pairs
40 TYPE: nucleic acid
41 STRANDEDNESS: double
42 TOPOLOGY: linear
43 MOLECULE TYPE: DNA (genomic)
44 FEATURE:
45 NAME/KEY: CDS
46 LOCATION: 144..6857
47 FEATURE:
48 NAME/KEY: 5'UTR
49 LOCATION: 1..143
50 FEATURE:
51 NAME/KEY: 3'UTR
52 LOCATION: 6855..7175
53 IS-08-450-562-8

```

Db	2928	GAGGGGAGCCCCAGCCACACCGGGCACCGGCACACAGAGTCCGAGCAAGGATGTGCC	2987
Qy	1095	gaagcgcgtggtgtgtcgtgacaggagggcgcaattttctcgtcagaagaagaaagtgcg-g	1153
Db	2988	GGCCCCAAGGGCCGAGCGCGCGCGCGGCGGACCGCGGCGGCCCCGAGCGGGGCCGGGAG	3047
Qy	1154	gcgctgagcccccagagagagggcgacgttgtatagccagtgtaaccgttgacgcttcgt	1213
Db	3048	GGGAGAGCGGGGGAGGAGCGCGGGCGGCGGCACCGGGCCCCGGCACAAAGCGCAGCTTCT	3107
Qy	1214	gctgagagacaaagaagaagtgtccagagcttggaaggcgatgtatgtagaaagcaaaagaagaa	1273
Db	3108	-CACGAGAGCTGTGTGAGAGAGCGACCCACGGAGAAAGGAGCGCCACGAGGAAGAGAGCTGAGAT	3166
Qy	1274	ggtgaaagcgaggttgaaatgagttc	1301
Db	3167	AGTGGAAAGCCGACAAAGAAAGAGCTTC	3194

```

      11
US-08-713-118-1
: Sequence 1, Application US/08713118
: Patent No. 6040436
: GENERAL INFORMATION:
: APPLICANT: Franco, Rodrigo
: APPLICANT: Sun Chen, Ai Pu
: APPLICANT: Sney, David J.
: TITLE OF INVENTION: NUCLEIC ACID ENCODING HUMAN NEURONAL
: TITLE OF INVENTION: CALCIUM CHANNEL SUBUNITS
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
: STREET: Two Millitia Drive
: CITY: Lexington
: STATE: MA
: COUNTRY: USA
: ZIP: 02173-4799
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/713.118
: FILING DATE: 16-SEP-1996
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Mata, Elizabeth W.
: REGISTRATION NUMBER: 38,236
: REFERENCE/DOCKET NUMBER: ACC96-01
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-861-9540
: TELEFAX: 617-861-6240
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7266 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 92..7102
: US-08-713-118-1

Query Match      2.9%, Score 46.6; DB 3; Length 7266;
Best Local Similarity 46.7%; Pred. No. 0.16;
Matches 209; Conservative 0; Mismatches 237; Indels 2; Gaps 2.

```


Db 3108 -CACGAGGCTGTGGAGAGAGACACGAGAGAGAGGCCACGAGAGAGAGGCTGAGAT 3166
Oy 1274 ggtggaaggcgaggtggaatggaagtc 1301
||||| ||| ||||| |||
Db 3167 AGTGAAGCCGACAAAGGAAAAGGAGCCTC 3194

Search completed: June 28, 2001, 20:25:17
Job time: 5605 sec

117: gb_est48:*
 118: gb_est49:*
 119: gb_est50:*
 120: gb_est51:*
 121: gb_est52:*
 122: gb_est53:*
 123: gb_est54:*
 124: gb_est55:*
 125: gb_est56:*
 126: gb_est57:*
 127: gb_est58:*
 128: gb_est59:*
 129: gb_est60:*
 130: gb_est61:*
 131: gb_est62:*
 132: gb_est63:*
 133: gb_est64:*
 134: gb_est65:*
 135: gb_est66:*
 136: gb_est67:*
 137: gb_est68:*
 138: gb_est69:*
 139: gb_est70:*
 140: gb_est71:*
 141: gb_est72:*
 142: gb_est73:*
 143: gb_est74:*
 144: gb_est75:*
 145: gb_est76:*
 146: gb_est77:*
 147: gb_est78:*
 148: gb_est79:*
 149: gb_est80:*
 150: gb_est81:*
 151: gb_est82:*
 152: gb_est83:*
 153: gb_est84:*
 154: gb_est85:*
 155: gb_est86:*
 156: gb_est87:*
 157: gb_est88:*
 158: gb_est89:*
 159: gb_est90:*
 160: gb_est91:*
 161: gb_est92:*
 162: gb_est93:*
 163: gb_est94:*
 164: gb_est95:*
 165: gb_est96:*
 166: gb_est97:*
 167: gb_est98:*
 168: gb_est99:*
 169: gb_est100:*
 170: gb_est101:*
 171: gb_est102:*
 172: gb_est103:*
 173: gb_est104:*
 174: gb_est105:*
 175: gb_est106:*
 176: gb_est107:*
 177: gb_est108:*
 178: gb_est109:*
 179: gb_est110:*
 180: gb_est111:*
 181: gb_est112:*
 182: gb_est113:*
 183: gb_est114:*
 184: gb_est115:*
 185: gb_est116:*
 186: gb_est117:*
 187: gb_est118:*
 188: gb_est119:*
 189: gb_est120:*

190: gb_est121:*
 191: gb_est122:*
 192: gb_est123:*
 193: gb_est124:*
 194: gb_est125:*
 195: gb_est126:*
 196: gb_est127:*
 197: gb_est128:*
 198: gb_est129:*
 199: gb_est130:*
 200: gb_est131:*
 201: gb_est132:*
 202: gb_est133:*
 203: gb_est134:*
 204: gb_est135:*
 205: gb_est136:*
 206: gb_est137:*
 207: gb_est138:*
 208: gb_est139:*
 209: gb_est140:*
 210: gb_est141:*
 211: gb_est142:*
 212: gb_est143:*
 213: gb_est144:*
 214: gb_est145:*
 215: gb_est146:*
 216: gb_est147:*
 217: gb_est148:*
 218: gb_est149:*
 219: gb_est150:*
 220: gb_est151:*
 221: gb_est152:*
 222: gb_est153:*
 223: gb_est154:*
 224: gb_est155:*
 225: gb_est156:*
 226: gb_est157:*
 227: gb_est158:*
 228: gb_est159:*
 229: gb_est160:*
 230: gb_est161:*
 231: gb_est162:*
 232: gb_est163:*
 233: gb_est164:*
 234: gb_est165:*
 235: gb_est166:*
 236: gb_est167:*
 237: gb_est168:*
 238: gb_est169:*
 239: gb_est170:*
 240: gb_est171:*
 241: gb_est172:*
 242: gb_est173:*
 243: gb_est174:*
 244: gb_est175:*
 245: gb_est176:*
 246: gb_est177:*
 247: gb_est178:*
 248: gb_est179:*
 249: gb_est180:*
 250: gb_est181:*
 251: gb_est182:*
 252: gb_est183:*
 253: gb_est184:*
 254: gb_est185:*
 255: gb_est186:*
 256: gb_est187:*
 257: gb_est188:*
 258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

JOURNAL REFERENCE AUTHORS	Methods Enzymol. 303, 19-44 (1999)
TITLE	2 (sites)
JOURNAL MEDLINE REFERENCE AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
TITLE	3 (sites)
JOURNAL MEDLINE REFERENCE AUTHORS	Shibata, S., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Kikuchi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaueuchi, S., Ikegami, T., Keshavag, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Okazaki, Y., Muramatsu, M., Inoue, Y. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
TITLE	4 (sites)
JOURNAL MEDLINE REFERENCE AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
TITLE	5 (bases 1 to 2374)
JOURNAL REFERENCE AUTHORS	Aachchi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hasegawa, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, T., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (18-AUG-2000) yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute: 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/ , Tel: 81-45-503-9222, Fax: 81-45-503-9216)
TITLE	6 (sites)
JOURNAL REFERENCE AUTHORS	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to

```

CDS
/dev_stage="0 day neonate"
101..2059
/note="putative"
/codon_start=1
/protein_id="BAB31756.1"
/db_xref="GI:1285731"
/translation="MDSGRVWAGAEMLIRVSPQVLTLDIPVSKKEELSYPLDAIVRCIE
AVMPGRSRLSLILLCODEPERAOPDVHFEQQLLGAELIRDDIGALONVSGSGGER
AAARATQERISALVAPAEETPLOLRSPVGLVINTVEPSAVRRPQVESLPETREEAR
KPLDARTTSSADPTSPDGLPPRGELIAGLRDVIDLHNHVDVESFVSRLOKSAEATM
RVLHRRGRTGRTTARRAAGEGLLTLLRAKPTPAETAYDVLQIKVAFSLARLQINAP
SSPELLHLEFGPLQMIIVTSGSGPEFAKSVRRPHLTLEVTLLRNNVPGENELTSLG
DSMTCPGVELPEEGESPYSPFTYNGMEPPATDPGGRPEDEVEERLOHEKRRQOASAP
QVAAHGOODEGPELTETESOLEEKARKVWVLCNFOFARNSELTSKRIADVAEYDDEKRR
KVRRHGOEGVPEYNNILPLHPGPOVHRSSQSPARLLETTPPEARVAPAPVYRQMD
SCSDSLNDLSEKEKFSQMLCVNELIORSIAAGSGSGRNTYRGRAABEQLSPSEMS
VVRAMLQTKGSSCTVRLALVLTTCALPFLSLOKEELRAVCPDEGARVTSQYTVQALLLE
DREYVSELEAEMEROKKVEGETEYV"
BASE COUNT      555 a      684 c      722 g      413 t
ORIGIN

```

Query Match	51.0%;	Score 833.8;	DB 192;	Length 2374;
Best Local Similarity	80.6%;	Prote. No. 1.9e-173;		
Matches 1054;	Conservative	0;	Mismatches 230;	Indels 24; Gaps 6;

QY	3	gcccgcagaggaagccgcggcgccggtctgggaaagctctctaaacgtctggggcaaac	62
Db	777	GAGGCCAGGACACAGACGCCGAGCTGCGGGAGAGAGGCTCTGCTGCACTGAGGGCCAAAGC	836
QY	63	cgccctcgagagccgagltacacacgactgctctcgagaagatcaagtacgctctcacctgc	122
Db	837	CACCCACTGAGGCGTGAGTACACCGACGCTGCATGAGAAATCAAAATACCCCTTCACGTGCG	896
QY	123	tggcccgagctgcgcggcaacaatcgccgagccctctcccgagagtgtgtgaattcctt	182
Db	897	TGGCCCGGTTGCGAGGCAACATCGCCAAACCCCTCTCTCCAGAGGTGTGTGACATTTCTGT	956
QY	183	tccggcctctgcagaatgtgtgaaacagctgtggggggccggagttctcgacagcagtgc	242
Db	957	TTGACCCCTGCAATGATGTGTGAACACTTCAGGGCGCCCTGAGTGGCGGAAGATGTC	1016
QY	243	gggggcccgcattcaatcgaatgcggtgcggtgcgctgcgcgagaaacagtaactccagtg	302
Db	1017	GACACCGCATTTTGACAGTGTGAGGCTGTGATCTGTCTCCGGACAAATGTCACCTCGTGTG	1076
QY	303	aaacagagctcttgaaactcgtctgtggggaactcgttgaaccgcccgcgggtgtgaagctgcc	362
Db	1077	AAATGAACTATGGAACCTCGCTGGGAGACACTCTGAGACTGCCCCAGGGGTGAGCTGCCCC	1136
QY	363	cggagggaggagccccccatcacagacccgaagtctctcaagcggctcttggagagccgcggtactag	422
Db	1137	CGGAGAGAGGGGTCTCCATACAGTCCCGAATTCACAAATGCGTGGAGAGCTCCGCCCACTG	1196
QY	423	accgcagagagccgcgctctgtagaagaccagttgagaacaaactacagacagacggaaagc	482
Db	1197	ATTCACAGGGCGCGCCCTGCGGAGGACCCAGTGGGAAAAACACTACACACGACGAAGCGCG	1256
QY	483	gcccgcagcaaaagcgcgcccgaggtgcgtctgtaaatgtgtcacaggaacttggagccaaga	542
Db	1257	GCAAGCGACGAAGAGGCCGCCCGAGGTGCGGGTCAATGTGTACACAAG--ATCCAGGAACCTGGA	1314
QY	543	atctgaagcctcagcttgtagtctagaagaacagcaggaataaggtctcgtgttaattgaact	602
Db	1315	AACGTGATCCACGCTAGA--AGAGAAAGCCAGAAATGGGTTTGTGTATTATTGACTT	1371
QY	603	ccaagcccgcaaacagctgaagctgtcgtgtcaagcagcggagcttaactggaagctctgga	662
Db	1372	CCAGGCCCGCAAAATGGCAGCGAGCTATGTGTCACACCGAGATGTGTTGGAGGTCGTGGA	1431
QY	663	tgaagatcgttaagtggtgggaaggtctggggcccgccggggcggaagaagatatgtgcccta	722
Db	1432	TGAGAGGCGCAAAATGGTGAAGGTTCGAGACCATTCAGGAGCAAGAGGGCTATGTGACCTTA	1491

QY	723	caaatctctcaacacctaccccgagcccgagctggtctacacagaccaaagccctgccgcag	782
Db	1492	TAAATTCGTACACCCACCCTCGAGCTCTAGGTGCATCCGACCCAAACCCCGAAGACA	1551
QY	783	ccctgaacagactctctctccacacagccccagcccccgccccacactcagctctgcg	842
Db	1552	CTGTGAAGACTAGACACACTCTCCGCCA---CCGTGCCAGCTCCAGGCCCACTAGGT	1608
QY	843	tctggcccgctgggaagagcccgcttgggaacgctgcgatnctcaacgcttggacccc	902
Db	1609	GCGGCCCAATGGAGACAGTTGGACAAG-----TCTAAACAGTTTGGATCCC	1654
QY	903	agcgaagaagaagaatctccacagatctcatctgcacagagaagaacagcagcgcgctg	962
Db	1655	ACGGAGAGAGAGAAATTCCTCCAGATGCTTTGTGTGCAATGAGAGACTGCACTCCGCGCTG	1714
QY	963	gccccagggccgctcgggaacgagccgcgcgaatcccaagggccccgcgccccggaacgcag	1022
Db	1715	GGCAGAGGGCGGTTCGGGTGCCAGCGGGGTAACCCCGGGGCCCGCAGGCCAAGACCCCAAG	1774
QY	1023	ctgaagccggagctcgggaacgctccgamtctcgcgctctgctgcgaagmcaagggcttagc	1082
Db	1775	CTCAGCCCGGGCTCTGAGGCGCTGGTGTGTCGTGCTGCTGTCGACAGCAAGGCGTTTATGC	1834
QY	1083	tccggagacgctgagcgcgctgtgtgtctgtacagcggggcgca-cttctctcgctcagaag	1141
Db	1835	TCCGGGACACTGTGAGAGCGCGCTCGTGTGCTGACACCGCGCACAGCTCTTCTCGCTGCAAAAG	1894
QY	1142	gaagagctgcgggacgtgtagcccccagagaggggacgctgtgtacagccagtcaccgtg	1201
Db	1895	GAAAGATTCCGGGGCGGTGTGCCCCACGAGNAAGGGCGCGCAGTGTACACCCAAGTCACGTGTG	1954
QY	1202	cagcgcttcgctgcgtcggagagacaagaagatgttccagagcttggagcagctgattgagaa	1261
Db	1955	CACGCG- GCGGCTGCTGGAAGACAGAGAAAAGTATCAAGCTGGAAGCGCTGTATGAGAA	2013
QY	1262	gcacaaagaagaagtgtagaagcgagagtgtagtaagtctatttgcac	1309
Db	2014	GCAAAAGAAAAGAAAGTGGAAAGCGACCAACCAAAACAGAACTCATTTTGATC	2061
RESULT	2		
LOCUS	AK014556		
DEFINITION	Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched library, clone:4632407K17, full insert sequence.		
ACCESSION	AK014556		
VERSION	AK014556.1	GI:12852484	
KEYWORDS	CAP trapper.		
SOURCE	Mus musculus (strain:C57BL/6J) 0 day neonate skin cDNA to mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library clone:4632407K17.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (sites)		
TITLE	Carninci.P. and Hayashizaki.Y.		
JOURNAL	High-efficiency full-length cDNA cloning		
REFERENCE	Methods Enzymol. 303, 19-44 (1999)		
AUTHORS	2 (sites)		
TITLE	Carninci.P., Shibata.Y., Hayatsu.N., Sugahara.Y., Shibata.K., Itoh.M., Kono.H., Okazaki.Y., Muramatsu.M. and Hayashizaki.Y.		
AUTHORS	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
TITLE	Genome Res. 10 (10), 1617-1630 (2000)		
JOURNAL	20493374		
MEDLINE	3 (sites)		
REFERENCE	Kono.H., Akiyama.Y., Nishi.K., Kitsuai.T., Tashiro.H., Itoh.M., Kikuchi.N., Ishii.Y., Nakamura.S., Hazama.M., Nishine.T., Harada.A., Yamamoto.R., Matsumoto.H., Sakaguchi.S., Ikegami.T., Kasaihagi.K., Fujiwara.K., Inoue.K., Togawa.Y., Izawa.M., Ohara.E.		
AUTHORS			


```

Db 2141 AAGGAAGACTTGGCGGCTGTGCCCCCGAGAGAGGGCGCCGAGTGTACAGCCAGTCACT 2200
Oy 1199 gtgcagcgcttcgtctgtgagagacaagaagtgatcagagctgagagcgatgata 1258
Db 2201 gtgcagcgcttcgtctgtgagagacaagaagtgatcagagctgagagcgatgata 2259
Oy 1259 gaagcaaaaagaagagtgagagcgagtgagagtgagagtgagagtgagagtgagagc 1309
Db 2260 GAAGCAAAAAGAAAAGTGAAGGCGAGACCAAAACAGAAATGATTTGATC 2310

RESULT 3
LOCUS AM972535 616 bp mRNA EST 01-JUN-2000
DEFINITION EST384626 MAGE resequences, MAGL Homo sapiens cDNA, mRNA sequence.
ACCESSION AM972535
VERSION AM972535.1 GI:8162381
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 616)
Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C., Holt,
I. E., Saeed, A. I., Sharov, V., Lee, N. H., Yeatman, T. J. and
Quackenbush, J.
Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
Unpublished (2000)
JOURNAL Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 307
Seq primer: Forward.
FEATURES
Source Location/Qualifiers
1..616 /organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGL"
/note="Vector: pBluescriptSkm"
BASE COUNT 128 a 205 c 192 g 89 t 2 others
ORIGIN
Query Match 35.5%; Score 580; DB 122; Length 616;
Best Local Similarity 97.7%; Pred. No. 1.3e-117;
Matches 598; Conservative 0; Mismatches 12; Indels 2; Gaps 1;

Oy 244 ggcgcgcgcatcgtgacatgcgacgctgagcgtgctgcgagacaacgctcactccagctga 303
Db 7 CGGCGCCGATCGACATCGATGCGCGCTGCGCGCTGCGGAGCAACGTCACCTCCACGTGA 66
Oy 304 aagaagagctcgtgacgtcgtgagagagctcgtgagaccccgcccgagctgagagctcccc 363
Db 67 AAACGAGCTCTGGAGCTCTGCTGGGGGACTCGTGGACCCCGCGGCTGGAGCTGTCTCCC 126
Oy 364 ggaagagagaccccatcacagaccgagcttctcagcgctgagagcgccgagctgactga 423
Db 127 GGAGAGAGGACCCCATACACACCGAGATTCTTCAGCGGCTGGAGCGCGCGGCTCAGTGA 186
Oy 424 ccgcagagagcgcgctgagagagacccagctgagagaaacagctacagcagcagagagcg 483
Db 187 CCCCAGAGCCGCGCTGAGAGACCCAGTTGAGAAACAGTACAGCAGCAGAGCGAGGCG 246
Oy 484 ccgcagcagaagcgcccccgagtgctgctgtaatggttaccgagagacttgagagcagaa 543
Db 247 CCGGCAAGCAAAAGCCGCCCGAGGTCGTCAATGGTCAACGAG--ACTTGGAGCCAGAA 304
Oy 544 tctgagcctcagctgagatcagagacagcagaaatggtcctggtgtaataatgacttc 603

```

```

Db 305 TCTGAGCCTCAGCTGAGTCAAGACAGACAGCAAAATGCTCTGTGTAATATGACTTTC 364
Oy 604 caggcgcgaacagcagctgagctgtcgtgtaacagcagcgagacgtactgagctctgagat 663
Db 365 CAGGCCCGCAACAGCAGTGAAGTGTGCGGTCAAGCAGCGGAGCTACTGAGAGTCTTGAT 424
Oy 664 gacagctgtaagtggtggaaggttcggagaccagcgagcgagagagagatgagcctac 723
Db 425 GACAGTCTGAAGTGTGGAAGGTTGGGAGACCAGCGGGGAGAGAGATATGTGCCCTTAC 484
Oy 724 aacatcctgaaacctaccaccgagcccgagctgacacacagcaaaccttgccgacg 783
Db 485 AACATCTGACACCTTACCCCGGACCCGCGCTGACACACAGCCAAAGCCTGCGCGACG 544
Oy 784 ctgaacagcactcctcctccaccacagcccgagcccgagccacactcagactgagct 843
Db 545 CTGAACACACTCTCCCTCCTCAACACACAGCCAGCCCGGCGCCACCTTCAAGCTTCTCT 604
Oy 844 cggcccgagctg 855
Db 605 CGGCGCCGCTGG 616

RESULT 4
LOCUS BE396214 743 bp mRNA EST 21-JUL-2000
DEFINITION 60131021F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3631850 5',
ACCESSION BE396214
VERSION BE396214.1 GI:9341579
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 743)
NIH-MGC http://mgc.ncl.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rgabs@emall.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNCM522 row: e column: 03
High quality sequence start: 44
High quality sequence stop: 691.
FEATURES
Source Location/Qualifiers
1..743 /organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:3631850"
/clone_lib="NIH_MGC_44"
/tissue_type="endometrium, adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; vector: pORF7; site_1: XhoI; site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAAGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT 159 a 242 c 233 g 109 t
ORIGIN
Query Match 35.1%; Score 573.4; DB 167; Length 743;
Best Local Similarity 94.3%; Pred. No. 3.7e-116;
Matches 660; Conservative 0; Mismatches 32; Indels 8; Gaps 6;

```


[illegible]

QY	658	CGGGGCAAGAGGATATgTgCCtCaCaCaatCtGyaCCtaccCCGgagCCCGgTgC	757
Db	359	CGGGGCAAGAGGATATgTgCCtCaCAAACTCTGACACCTCAGCCCGGAGCCCGGCTGC	418
QY	758	accacagccaagagccctgTgCCCGagccagctgtaaCagcaCtctcctcCaaccagcccc--	815
Db	419	ACCAACAGCAAAAGCCCTgTgCCCGCAGCCCTGAAACACACACTCTCTCCACACAGCCGCCCA	478
QY	816	-agccCCgTgCCCCaCtCaagTgctTgTgCtGgCCCGcTgTgagCaGgCCCGc- tggYac	873
Db	479	AGCAACCGGGCCCCACCTTCACACTCTgGCTCAGGCCCCGTGGAGACGGCACCGCATGGGAC	538
QY	874	agctgCgata-ncTcaacgTctTgTgacCCcagCGaagaagaaatctctccagatgctc	932
Db	539	AGCTGCGATACCTCTCAACGGGTAGACACCAAGAGAAAGAAACATCTTACAGATGCTC	598
QY	933	atcgttcaacagagatctcagagCGcctTgCCaagGgCCgctcTogagccGaaCCgCgca	992
Db	599	ATCGACAA-GAGGAATCGCAAGAGCGCTGCGCACAGGCCAGGAGGACAGCGACAAACCA	657
QY	993	gtccag 999	
Db	658	GGCCGAG 664	

RESULT	7
LOCUS	BE744353
DEFINITION	BE744353 937 bp mRNA EST 15-SEP-2000 601576580F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3837570 5' , mRNA sequence.
ACCESSION	BE744353
VERSION	BE744353.1 GI:10158345
KEYWORDS	EST .
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Makayota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 937) NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE	Unpublished (1999)
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: DCTD/DRP
COMMENT	cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL), DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov Plate: LLCW522 row: p column: 19 High quality sequence start: 4 High quality sequence stop: 671. Location/Qualifiers 1.937 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3837570" /clone.lib="NIH_MGC_9" /tissue.type="adenocarcinoma cell line" /lab.host="DH10B (phage-resistant)" /note="Organ: Ovary; Vector: pOTB7; Site_1: XhoI; Site_2 EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adapter: GCGACGAC(G). Size-selected >500bp for average insert size 1.kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)." .
BASE COUNT	177 a 297 c 313 g 150 t
ORIGIN	

Accession	BE543227
Version	BE543227.1
Keywords	GI:9771872
Source	EST.
Organism	human. Homo sapiens Eukaryota; Euteleostomi; Mammalia; Euteria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Primates; Catarrhini; Homidae; Homo.
Reference Authors	NIH-MGC
Title	http://mgc.ncl.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC)
Journal	Unpublished (1999)
Comment	Contact: Robert Strusberg, Ph.D. Email: cgaphs-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: Incyte Genomics, Inc. DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/JLNL at: http://image.llnl.gov Plate: LLM8441 row: p column: 05 High quality sequence stop: 587. Location/Qualifiers
FEATURES	1..596
source	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3455476" /clone_lib="NIH-MGC_12" /tissue_type="cervical carcinoma cell line" /lab_host="DH10B" /note="Orgn: cervix; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.4 kb. Library prepared by Life Technologies."
BASE COUNT	115 a 205 c 195 g 81 t
ORIGIN	
Query Match	30.5%; Score 498.6; DB 136; Length 596;
Best Local Similarity	96.0%; Pred. No. 1.le-99;
Matches 574; Conservative	0; Mismatches 17; Indels 7; Gaps 6
OY	582 ggctctgtgtaattatgaattccagggcccgaacagcgagtgcgtgtgtaagaacgagc 641
Db	1 ggtcctgtgtaattatgaattccagggcccgaacagcgagtgcgtgtgtaagaacgagc 60
OY	642 ggaacttgtagtgccttcgatgatcatctgtaagtcgtgtaagtttcggaccacgagg 701
Db	61 ggacctactgcagagctccttcgatgatcatctgtaagtcgtgtaagtttcggaccacgagg 120
OY	702 gcaagagagatatgtgccttacacaatcctgcacacctaacccccgagaccggctgcacca 761
Db	121 gcagagagagatatgtgccttacacaatcctgcacacctaacccccgagaccggctgcacca 180
OY	762 cagccaagaagcctgcgcgcgagcctctaacaagcacctccttcacaccaagaagcccgccc 821
Db	181 cagccaagaagcctgcgcgcgagcctctaacaagcacctccttcacaccaagaagcccgccc 240
OY	822 gggcccaactcgaagctcgcgcctcggcccccgcctggagacagggcccgctggaaagttgga 881
Db	241 gg-cccaccttcgaagctcgcgcctcggcccccgcctggagacagggcccgctggaaagttgga 299
OY	882 ta-nctcaacaggtctgtaaccacgaagagagaatcttcacagatgctcatgctcaa 940
Db	300 TAGCTCAACGGGCTTGGAAGCCCGAGGAGAAGAAATTCTCCGAGATGCTCATGTGTA 359
OY	941 cgaaggaactgcagagcgc-gcctgcgccagggcgctcggagacgagcgcgcagc-agt-ccca 998
Db	360 CGAGGAATTCGACAGGGCAGCGCTGGGCCCGCTGGGACAG6CCCCCGCTGGGAAGCTGGCCA 419
OY	999 gggcccgccgcccggaaacgcgaagtcgaagcccggtctggagcgtctccgangtccgc--g 1056
Db	420 GGCCCCAAGCCCCCGAACCAGATGTCAGCCCAGGCGCTGGAGCGCTCCGAGGTCGCGACG 479

Qy	1057	ctcgagctcgaagcgaagggcgttaagcttcgcggagccgttgagcgcgtggtgtctacgg	1116
Db	480	ctggcctgcagacccaagggcttttagctccggagccctggagcgcgttgggtcttagccg	539
Qy	1117	ggagcgaccttttcgcgtcgcagaaggagagcgtgcggcggtgtagcccccagaagaggg	1174
Db	540	ggcgcgagtttactccctgcagagggagagcgtg-gggcgcgtgacgccccaagaaagg	596
RESULT	13		
LOCUS	Al183328/c		
DEFINITION	Al183328	542 bp	mRNA
ACCESSION	Al183328		EST
VERSION	Al183328.1		28-OCT-1998
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 542)		
TITLE	NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.		
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
COMMENT	Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.lnl.gov) for further information. Insert length: 898 Std Error: 0.00 Seq primer: -40up from Gibco High quality sequence stop: 276. Location/Qualifiers 1..542 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:1723555" /clone_l1b="Soares-placenta_8tc9weeks_2NBHP8tc9w" /dev_stage="two placentae: one from 8 weeks and another from 9 weeks post conception" /lab_host="DH10B (ampicillin resistant)" /note="Organ: placenta; Vector: pRT73D (pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCCGATGTTTTTTTTTTTTTTT 3']; double-stranded cDNA was size selected, ligated to Eco RI adapters (pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pRT73 vector (pharmacia). library constructed by Bento Soares and M.Fatima Bonaldo."		
BASE COUNT	93 a	191 c	136 g
ORIGIN		117 t	5 others
Query Match	30.0%	Score 489.8:	DB 17; Length 542;
Best Local Similarity	94.8%	Pred. No. 9.1e-98;	
Matches 514; Conservative	0;	Mismatches 27;	Indels 1; Gaps 1;
Qy	1074	ggccttaagctccggagaccgtgaagcgcgtgggtgtgctgacccggggcgca-ctttctcg	1132
Db	542	ggccttagctcggggagccctgagccgcnntggtgtgctgacggggggcgcagctttttctcg	483
Qy	1133	ctgcagaagagagagctgcggcggtgtgaagccccaagagaggggcaagtggtacagccag	1192
Db	482	ctgcag	423
Qy	1193	gtcaccgtgcagcgcttcgtcgtgtgtaggagacaagaagagtgctaaagagcttgaggcagt	1252
Db	422	gtcaccgtgcagagcgttcgtcgttcgtgagagcaaaaagagaaagtgctcagagagctgagagcag	363

Oy 1253 gatgagaagcaaaagaagatggaagcgaggtggaatgagctcattgacctgc 1312
 |||||||
 Db 362 GATGGAAGCAAAAGAGAGTGAAGCGAGCTGGAATGAGGTGATTTGACCTGCG 303
 |||||||
 Oy 1313 cagggcccttcgcgaagaagtgagagggcccgctggaagaagagctcctaagctcgc 1372
 |||||||
 Db 302 CAGGGCCCTTCGCAAAAGTGAAGAGCGCCCGTGGAGAAAGGAGCTCTGACTCTGC 243
 |||||||
 Oy 1373 ccaatagaggaagtcgaactctctggaatgagcaatctgctccgagctctcgc 1432
 |||||||
 Db 242 CCAATAGCGCAAGTGCATCTTCTGAAGAGATGCAATCTGCTCCGCGCTTCTCCCG 183
 |||||||
 Oy 1433 catcccggtggaagacttaacgaltctgctgagctccctccggaagaagatctgacctg 1492
 |||||||
 Db 182 CATCCCGGTGGAAGACTTAACGATCTTGTGCTGAGTCCCTCCGAGAGAGATCTGAGCTG 123
 |||||||
 Oy 1453 gctggagatgagggagggcgctggaagcaatctcagaaagcgctagcaagcccgagaagg 1552
 |||||||
 Db 122 GCTGGAGTGGGAGGCGGTGGAGACAGTCTACGAAAGCGCTACAGACCCCGAGAGG 63
 |||||||
 Oy 1553 gtgcagtgagccctgagcatctgataatgagcccgagctataacagagctcgctc 1612
 |||||||
 Db 62 GTGCAGTGGAGCCCTGACGATTTGTAATATGCGCGCCAGCTTAACAGCCTCGCTCTT 3
 |||||||
 Oy 1613 ag 1614
 ||
 Db 2 AG 1
 ||
 RESULT 14
 BE393472 674 bp mRNA EST 21-JUL-2000
 LOCUS 601310143P1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3631498 5',
 DEFINITION mRNA sequence.
 ACCESSION BE393472
 VERSION BE393472.1 GI:9338752
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 674)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LCM321 row: f column: 11
 High quality sequence stop: 628.
 location/Qualifiers
 1..674
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3631498"
 /clone_lib="NIH_MGC_44"
 /tissue_type="endometrium, adenocarcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: uterus; Vector: pOT7; Site_1: XhoI; Site_2:
 EcoRI; CDNA made by oligo-dt priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-CDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 132 a 243 c 210 g 89 t

Query Match 29.9% Score 488.4; DB 167; Length 674;
 Best Local Similarity 95.3%; Pred. No. 1.9e-97;
 Matches 611; Conservative 0; Mismatches 17; Indels 13; Gaps 10;
 Oy 338 accgccccgagctgagactgtcccgagagagagaccacacacagagctc 397
 |||||||
 Db 1 ACCGCCCCGGGCTGAGACTGTCCCGAGAGAGAGACCCCGATACAGACCGAGTCTTC 60
 |||||||
 Oy 388 a-gcgctgtggaagcgccggtcactgaccgcgagagcgcgctggaagagaccagttga 456
 |||||||
 Db 61 AGCGCGCTGGGAGCGCGGCTACTGACCCGAGAGCGCGCTGGAGAGACCGAGTTGA 120
 |||||||
 Oy 457 gaaacagctacagcagagggagggcgccgagagaagagcgcccgaggtctgctga 516
 |||||||
 Db 121 GAACAGCTACAGCAGCAGCGGAGCGCGGAGAGAAAGCGCCCGAGGTGCTCA 180
 |||||||
 Oy 517 tggtaacgagacttggaagcagaatctgagctcagctgagctgagagagacagaga 576
 |||||||
 Db 181 TGCTACCGAG--ACTTGGAGCCAGAACTGAGCTTCACTGAGTCAAGACAGACAGA 238
 |||||||
 Oy 577 aaatggtctctgtgtaactatgactccaagccgagcaagagagtgagctgagtcagg 636
 |||||||
 Db 239 AAATGGTCTCTGTGTAATATGATGATCCAGCGCCGCAACAGCAGTGAAGCTGCTCAAG 298
 |||||||
 Oy 637 cagcgagagcgtactgagaggtcctgagatgacagtcgtaagtgtggaaggttcggagcca 696
 |||||||
 Db 299 CAGCGGAGAGTACTGAGGCTCTGATGATGACAGTCTGAAGTGTGAGAGTTGGAGACCA 358
 |||||||
 Oy 697 gcgggagagagagatgctgctccacacatcccgagacccctacccgagcccgctg 756
 |||||||
 Db 359 CGGGGAGAGAGAGATATGTGCTGCTACAAACATCTGACACCTTACCCGAGACCGGCTG 418
 |||||||
 Oy 757 caccacagcaaaagccctgcccagcgtgaaagcactctctccacacacagcccca 816
 |||||||
 Db 419 CACACAGCAAAAGCCCTGCGCCGAGCTGAGACAGACACTCTCTCCACACAGACCCCA 478
 |||||||
 Oy 817 gcccgagcccaacccctcagctgtgctgagcccgctggaagagcccg--gctggagaca 874
 |||||||
 Db 479 GCCCGGG--CCGACCTCCAGCTCTGCTGGCTGG--CCGCTGGAGACGCGCCCTGCTGGAGC 536
 |||||||
 Oy 875 gctggagata-nctcaagagcttga--ccccagcaggaaga--gaaatctccagatgagc 930
 |||||||
 Db 537 GCTGGATAGCTCCACAGGCTTGAGACCCCGAGAGAGAGAGAGACATTTCTCCCGAGATGC 596
 |||||||
 Oy 931 tcatcgaacagagagactgagcgagcgctgagccagggc 971
 |||||||
 Db 597 TCATCGTC-ACGAGGAACTGC-GCGCGCGCTGGCCGCGCC 635
 |||||||
 RESULT 15
 LOCUS A1818998 503 bp mRNA EST 21-DEC-1999
 DEFINITION w160f08.x1 NCI-CGAP Lu19 Homo sapiens cDNA clone IMAGE:2407239 3',
 similar to SW:EP58. HUMAN O12929 EPIDERMAL GROWTH FACTOR RECEPTOR
 KINASE SUBSTRATE EP58. ;, mRNA sequence.
 ACCESSION A1818998
 VERSION A1818998.1 GI:5438077
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 503)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.

